ESCAP MEETING NO. 4 - 01/12/00

AGENDA

There was no agenda developed or used for the January 12, 2000 meeting.

ESCAP MEETING NO. 4 - 01/12/00

HANDOUTS

ANALYSIS OF C.A.P.E. FINDINGS ON 1990 PES TECHNICAL ISSUES

Sally M. Obenski
U.S. Bureau of the Census
9 June 2000

This paper was prepared in conjunction with the report, *Accuracy and Coverage Evaluation: Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of the Census*, Kenneth Prewitt, U.S. Bureau of the Census, June 2000, available online at http://www.census.gov/Press-Release/www/feasibility.html. The author's address is Sally Obenski, PRED, FOB 2, Rm 1103, U.S. Bureau of the Census, Washington, DC 20233, Sally.M.Obenski@ccmail.census.gov.

June 9, 2000

ANALYSIS OF C.A.P.E. FINDINGS ON 1990 PES TECHNICAL ISSUES

Sally M. Obenski

Section I: Background

On July 15, 1991, Secretary of Commerce Robert A. Mosbacher announced his decision[1] that according to eight pre-established guidelines[2], the Census Bureau had failed to demonstrate that the 1991 Post Enumeration Survey (PES) adjusted counts improved the 1990 census counts at all geographic levels. He cited this conclusion as his basis for not adjusting the 1990 census counts. However, he requested that the Census Bureau continue analyzing the PES data to see if technical concerns could be overcome so that the population base for the intercensal population estimates could be adjusted. Consequently, the Census Bureau established the Committee on Adjustment of Postcensal Estimates (C.A.P.E.) that continued extensive research begun in 1990 in an attempt to solve technical concerns about adjustment identified by Secretary Mosbacher.

In August 1992, C.A.P.E. issued its report and concluded that, on average, an adjustment to the 1990 base at the national and state levels for use in intercensal estimates would lead to an improvement in accuracy in those estimates.[3] The conclusion was not unanimous, but the large majority of the Committee agreed with the finding. The C.A.P.E. did not find that adjustment improved accuracy at small geographic areas (generally less than 100,000). Additionally, while some of the technical concerns identified by Secretary Mosbacher were fully resolved, others were not. These technical concerns centered around types and levels of errors in the PES. In January 1993, Census Bureau Director Barbara E. Bryant announced her decision not to adjust the population base for the intercensal population estimates with the PES findings for reasons similar to Secretary Mosbacher's--improvement could not be seen at all geographic levels used by intercensal data. [4]

---

[1]56 Fed. Reg. 33582-33642 (July 22, 1991).

[2]55 Fed. Reg. 9838-9661 (March 15,1990).

[3]"Assessment of Accuracy of Adjusted Versus Unadjusted 1990 Census Base for Use in Intercensal Estimates," Report of the Committee on Adjustment of Postcensal Estimates, Bureau of the Census, Department of Commerce, August 7, 1992.

[4]58 Fed. Reg. 69-78 (January 4, 1993).

The purposes of this analysis are to (1) describe the technical issues addressed by C.A.P.E., (2) discuss major findings, and (3) provide issue statuses as applied to the Census 2000 Accuracy and Coverage Evaluation (A.C.E.) survey.[5]

Section II: Technical Issues Reviewed by C.A.P.E.

C.A.P.E. was to investigate potential census adjustment for intercensal population estimates. The issue facing the Committee was whether the potential error in the PES adjustment was at a sufficiently low level to recommend the inclusion of results into intercensal estimates. Estimating the census undercount with acceptably small error and, in turn, using that knowledge to improve the census counts for all levels of geography were two complex and difficult tasks. The PES relied on Dual System Estimation (DSE) methodology that can be summarized as follows:

- Selected a random sample of block clusters and listed the housing units;
- After the census enumeration, conducted interviews with the sampled people and determined their address on Census Day;
- Attempted to match the interviewed people to their enumerations in the census at their Census Day addresses, searching outside the block to neighboring blocks if necessary;
- Resolved cases where a match was not confirmed by attempting to obtain more information from the respondents;
- Used information from similar people to "impute" missing information;
- In the sample blocks, determined whether census enumerations were correct or instead erroneous because of duplication, errors in applying definitions, other respondent errors, or fabrications by census interviewers;
- Divided the entire country into *poststrata*, groupings of people by race, Hispanic origin, age, sex, tenure, and other predefined variables, who were expected to have similar undercount rates;
- Computed DSEs for each poststratum from estimated match rates, rates of erroneous enumerations, and census results;
- Compared the DSE to the census count for each poststratum to compute a coverage factor (the relative measure of how much the poststratum was over- or undercounted);

[5] A detailed discussion on how C.A.P.E. determined whether adjusted or unadjusted numbers were more accurate using loss functions is described in "Analysis of C.A.P.E. Findings on PES Accuracy at Varying Geographic Areas," Sally Obenski and Robert Fay, U.S. Census Bureau, June 9, 2000.

- Applied the coverage factor for each poststratum at every geographic level of interest; and
- Tabulated the adjusted census results.

When C.A.P.E. began discussing whether to adjust the base for intercensal estimates, it started by reviewing five technical issues raised by Secretary Mosbacher about whether to adjust the 1990 census. Each involved concerns about error in the PES/DSE system that affect the quality and reliability of the PES estimates.

First Issue: Could problems in the smoothing model be resolved?

The 1990 PES design included 1,392 different poststrata. Poststratification is the dividing up of the population into groups with similar capture probabilities, i.e., undercounts. They were formed according to pre-identified variables, such as age, sex, race, Hispanic origin, tenure, and other variables thought to be associated with differing capture probabilities. Each person in the PES universe fell into a unique poststratum. Because the PES sample was only about 165,000 housing units, some of these poststratum groupings had high sampling variances. To offset this variance, a statistical modeling technique called smoothing was used. Although the smoothing process was successful at reducing the variance, it introduced a number of complexities, including a concern about its lack of robustness. Models that are used in statistical methods, such as the smoothing model, must be robust. That is, they must not be sensitive to relatively small changes in assumptions.

C.A.P.E. determined that it would take a large, intense, and uncertain research effort to resolve concerns about smoothing, so the Committee decided to reduce the number of poststrata to 357. By doing so, each new poststratum would have enough sample size to support stable estimates (i.e., the estimates would not have very large sampling variance), therefore requiring no smoothing. By eliminating the need for smoothing, C.A.P.E. concluded that the new stratification system adequately dealt with that issue. However, any reduction in the number of individual poststratum could lead to an increase in heterogeneity, that is, people with differing capture probabilities could be grouped into the same poststratum. Therefore, C.A.P.E. decided to assess whether heterogeneity was problematic and those results are discussed below.

Status

Smoothing will not be used in Census 2000. In fact, because of Census Bureau concerns about the complexity and robustness of statistical models, only the synthetic model will be used to produce the official population estimates from the A.C.E. for geographic levels such as states, counties, tracts and blocks.

Second Issue: Can estimated biases be removed from PES estimates?

If it were possible, nonsampling error or bias in the PES estimates would be removed before any potential adjustment. Therefore, the Committee investigated whether it would be possible to increase the accuracy of the estimates by removing bias. The Census Bureau's 1991 model of total error in the PES was based on 13 evaluation poststrata. As a result of the change in poststratification, C.A.P.E. computed a revised total error model based on 10 evaluation poststrata that estimated the amount of bias present in the DSEs. At the national level, the bias could be removed. However, because of the very small samples used to estimate the biases and the difficulties of modeling, C.A.P.E. was reluctant to use modeling to distribute the bias sub-nationally so that they could be removed.

A partial solution was to examine block clusters that contributed the most to the PES estimate of undercount--influential observations. Census Bureau researchers conducted an extensive review of the 104 block clusters (of about 5,000 total block clusters in the survey) that significantly affected the variance of the estimates. Expert matchers tried to remove all matching error and therefore any bias in the survey estimates due to matching. This review reduced the estimated national undercount by 0.1 percent. However, the bias reduction only applied to the 104 influential blocks. It was also during this analysis that a computer processing error was discovered after the release of the 1991 official PES estimates.

Codes that were attached to cases in clerical processing were incorrectly processed in the computer. Specifically, a problem was discovered in the computer editing of erroneous enumerations. The intended procedure was that if a person did not reside in the sample block cluster on Census Day, the census record was to be treated as an erroneous enumeration. The procedure that was followed, however, was that this edit (i.e., given erroneous enumeration status) was only applied if the census record was matched. Otherwise, the census record was treated as a correct enumeration. Errors went in both directions (increasing and decreasing the estimated undercount), but the net result was to reduce the estimated national undercount by 0.4 percent. Consequently, the official 1990 census undercount estimate was revised by C.A.P.E. from about 2.1 to about 1.6 percent, or 4 million.

Status

To date, the Census Bureau has not identified an adequate method of removing bias from the estimates. To do so requires empirical data on the distribution of error over geographic levels that is very difficult, if not impossible, to obtain. Consequently, the best option to minimize bias is to ensure that the A.C.E. is well-designed and executed. The Census Bureau has made a number of design improvements intended to reduce or contain bias levels from 1990. These include an improved poststratification scheme (discussed below), matching system, and automated systems.

The Census 2000 A.C.E. design includes several improvements to the matching processes. First, the matching system is fully automated, rather than manual as in 1990. It has a number of built-in edits and quality checks to reduce errors. Automating processes that have been analyzed and refined over a twenty-year period will make searching and matching easier and more reliable. Second, the matching processes have been centralized in one site, rather than decentralized at several as in 1990, allowing for more effective control--a single, well-trained staff will perform all matching at a single location. Third, a change in the treatment of people who have moved since Census Day (i.e., movers) will simplify matching. Unlike in 1990, it will be necessary to match only the people who resided in the sample area on April 1. Given that matching error was not a serious concern in 1990 and that matching has been improved for 2000, matching should not generate significant error in the 2000 Census A.C.E.

To increase efficiency and data quality, the Census Bureau is using laptops to conduct Computer Assisted Person Interviewing (CAPI) rather than the pen and pencil approach used in 1990. The use of CAPI in the 1998 dress rehearsal demonstrated improvements from these changes. To mitigate the risk of a computer processing error in 2000, the Census Bureau has adopted a set of best practices used in system and software development to improve quality. The following provide illustrative examples:

! To reduce risk, the Census Bureau has included software validation and verification strategies, such as independent software development of key computer programs;

! To reduce ambiguity and increase communication, the Census Bureau has adopted an improved documentation approach for technical issues;

! The Census Bureau has developed a Sample Design Control System to control, monitor, and validate the different phases of sampling, and;

! To validate the accuracy of the estimation processing streams, the Census Bureau is developing an Integrated Review System. This system should facilitate an understanding of data sources, files, outputs, and assist in verifying the accuracy of files.

Such improvements should result in a controlled, robust, and reliable computer processing environment. In summary, the Census Bureau has made some important changes to the A.C.E. design that build upon the C.A.P.E. poststratification design and other analyses of the 1990 PES that should reduce, or at least contain, the level of bias in the A.C.E. estimates.

Third Issue--Part A: Is the total error model complete?

The third major concern that C.A.P.E. addressed was whether the total error model used to evaluate the 1991 PES contained all components of error and whether the components were adequately measured. As a result of additional analysis, C.A.P.E. identified two more sources of error to be added to the 1991 model---late-late returns and out-of-scope cases. Subsequently, the Committee was confident that all components of error had been listed and considered. However, the Committee could come to no agreement about whether the level of measured component error was adequate. While agreeing that the evaluation sample sizes were too small to be reliable for several estimates of bias, the Committee concluded that, at this point, nothing could be done to improve their accuracy. Because analysis (called loss function analysis) to determine whether the census numbers or adjusted estimates were more accurate was dependent on the levels of estimated bias, the general conclusion was to use caution in evaluating the results.

Status

For Census 2000, the Census Bureau plans to employ a total error model to measure individual error components. Building on the experiences and concerns with 1990 evaluations and with the 1991 and 1992 error models, problems experienced in 1991 should not be repeated.

Third Issue--Part B: Correlation Bias

As stated earlier, in addition to measurement biases such as matching error, recall bias, and fabrication, the DSE also contained correlation bias, which can occur if any of the following assumptions are not met.

- A person's participation in the PES is not affected by his or her participation in the census (the causal independence assumption). A breech of this assumption leads to contamination.

- Within each poststratum, persons have the same probability of inclusion in at least one of the two systems, the census or the PES (the homogeneity or synthetic assumption). This second component occurs because of variable capture probabilities within a poststratum. A breech of this assumption leads to heterogeneity.

- A special case of a failure of the homogeneity assumption is the set of people with zero probability of being captured--those missed in both the census and the PES--sometimes known as the unreachable people.

Contamination was not considered a big problem so C.A.P.E.'s analysis focus was on the failure of the homogeneity assumption. The Committee obtained an estimate for correlation bias by comparing PES estimates to Demographic Analysis (DA) estimates. DA is a well-developed coverage measurement and evaluation program where analytic estimates of net undercount are derived by comparing aggregate sets of data, such as birth, death, immigration, and Medicare statistics. As part of 1990 evaluations, PES estimates were compared to DA estimates. It was generally assumed that the PES understated the undercoverage of some populations as compared to DA. Generally, the DA estimates of males were higher than the PES. These added males were an estimate of the level of correlation bias in the PES. This estimate was added to the total error model and used in loss function analysis to determine whether the census numbers or adjusted estimates were more accurate. The Committee concerns included:

- It was not possible to disaggregate the people with variable capture probabilities from the unreachable people;

- The method used for comparing the DSE bias to DA understated the estimate of people missed due to correlation bias, and;

- The method of allocating the correlation bias to sub-national levels was uncertain: The estimated missing people (all males) were allocated back to each poststratum proportional to the estimate of the number of males estimated to be missed in both the PES and the census for the poststratum. Further modeling was used to allocate the total error down to sub-poststratum levels.

The Committee concluded that correlation bias should be a component of total error. However, because of the concerns discussed above, the Committee requested that loss function analysis be conducted with and without correlation bias.

Status

Although the PES and A.C.E. surveys are not perfect--they do not capture every person missed by the census--the DSEs do generally move the count *closer* to truth. For example, in 1990, DA estimates indicated that about 1.3 million Black males were missed by the census; the PES found about 730,000 of them. Further, it does not seem unreasonable to assume that people who are missed by both the census and the survey are distributed similarly to the known undercounted groups. Finally, although the model used to evaluate the performance of the adjusted numbers may not perfectly allocate the correlation bias sub-nationally, again, the Census Bureau designed the model to allocate the unreachable people similarly to the known undercounted.

Generally, correlation bias is expected to result in underestimation of the true population by the DSEs. However, since DSEs themselves generally exceed the corresponding census counts due to census undercoverage, the unadjusted census counts are subject to even larger downward biases than the DSEs. That is, the bias in the census counts includes *both* census undercoverage relative to DSEs, and correlation bias of the DSEs. Until the Census Bureau is able to gather more empirical data on sub-national levels of correlation bias, it is not prudent to consider including those people missed because of correlation bias in the official estimates of population. However, for evaluation purposes, the Census Bureau is investigating several approaches to obtaining better data and improving how bias is allocated.

For example, senior statisticians are examining whether it is efficacious to include both male and female DA estimates for adjusting the target population counts used in the evaluation of the census and the adjusted counts.

Third Issue--Part C: Loss Function Analysis

As mentioned earlier, a detailed discussion of the C.A.P.E. loss function work is covered in another analysis. However, in planning how to conduct the loss function analysis, the Committee could not come to consensus on:

- The best way to allocate the bias to produce target population numbers and to sub-poststratum levels of geography;
- The appropriate comparison among census, adjusted, and target numbers (e.g., simple difference, square of difference, absolute difference); and,
- Whether to include correlation bias.

Consequently, five forms of loss functions were run using four models of bias with and without correlation bias on numerous levels of geography, e.g., states, counties smaller than 200,000, places less than 25,000. What the Committee did agree on was that aggregate losses should be examined. That is, they agreed to look at the aggregate loss over all areas of interest (e.g., states) rather than individual losses. In examining aggregate loss over a set of areas, C.A.P.E. conducted statistical hypothesis tests on the loss function analyses results to ensure that the difference in aggregate loss between the census and the adjusted counts was a real difference rather than random error. Additionally, C.A.P.E. examined the distributive accuracy of a number of place sizes using a number of comparisons. In general, improvements to distributive accuracy were only demonstrated when places that were compared had different levels of undercount. For example, comparing large cities that all had similar undercounts did not demonstrate that adjustment improved distributive accuracy. However, when the large cities were compared to the balance of the nation, improvement was demonstrated. Generally, including correlation bias favored adjustment while excluding it did not, but the Committee concluded that correlation bias should be included in comparisons of accuracy.

The C.A.P.E. analyses indicated that adjustment improved the distributive accuracy of states and large counties and cities. Additionally, as different comparisons were made, improvements were beginning to be discerned in smaller areas even in cases where correlation bias was not included.[6]

---

[6] "Additional Research on Accuracy of Adjusted Versus Unadjusted 1990 Census Base for Use in Intercensal Estimates, Addendum to Report of the Committee on Adjustment of Postcensal Estimates," Bureau of the Census, November 25, 1992.

Status

Although it may not be measured and allocated perfectly, correlation bias is a proven phenomenon that needs to be taken into account. Consequently, correlation bias should be included in any assessment of relative accuracy between unadjusted census and adjusted counts. Since the A.C.E. design was based on and improved the C.A.P.E. PES design, if the 2000 census has an undercount comparable to 1990, distributive accuracy should be improved, on average, in states, large counties and cities, and be at least as good as the census enumeration in other areas.

Fourth Issue: Does the homogeneity assumption hold?

As part of the 1990 evaluations of the PES, the Census Bureau designed an evaluation study to assess whether the homogeneity assumption held, but the results were mixed or inconclusive. Consequently, given its importance and the reduction in poststrata, the C.A.P.E. asked for new research called artificial population analysis on whether the homogeneity assumption held. An integral part of the PES/DSE system is to assume that every person within a poststratum has approximately the same capture probability. Failure of this assumption leads to bias in the DSE. This assumption underlay the PES poststratification design and was used to make estimates for states and cities. Only the poststrata had direct estimates from the PES. Therefore, an indirect or synthetic adjustment assumed that the probability of being missed by the census was constant for each person within an age, race, Hispanic origin, sex, and tenure category in a geographic area within each poststrata.

C.A.P.E.'s concern was the degree to which the homogeneity assumption held--how much did the PES results differ from truth and what was the effect of that difference on PES accuracy? In brief, artificial population analysis was an attempt to see if the new poststratification was reliable by substituting new--surrogate--variables that were believed to be correlated with undercount, e.g., mail return rate of census questionnaires. The actual values of the surrogate variables were known (as opposed to the target population counts that were estimated). The loss functions were run and, if the loss was small, one could assume that the poststratification was reliable and the homogeneity assumption was holding. In addition, C.A.P.E. examined the number of places improved by adjustment since it knew the truth (the actual value of the surrogate variable).

For states and most large geographic areas, without any bias, artificial population analysis supported the homogeneity assumption. Once bias was introduced, however, the artificial population analysis showed less and less homogeneity. When bias reached 25 percent of the estimate, the artificial population analysis indicated a serious loss of homogeneity.

At the time of the August report, the Committee could only support the homogeneity assumption with some concern since the level of bias in the PES (22 percent when correlation bias was included) was close to the point where artificial population analysis showed the homogeneity assumption failing.

Residual heterogeneity, that is, heterogeneity in census inclusion probabilities not explained by the poststratification or model used to estimate census coverage, has two undesirable consequences. At the poststratum level and above the consequence of concern is correlation bias. For areas below the poststratum level, the concern is that geographic variation in inclusion probabilities violates the synthetic assumption, leading to biased estimates for small areas. For C.A.P.E., the primary concern was the former because intercensal estimates were not computed for very small areas.

Subsequently, a problem was identified with the study's methodology. Although most of the C.A.P.E. analyses were all based on distributive accuracy, the statistical analyses leading to the figure of 25 percent bias were all based on numeric accuracy. Therefore, the artificial population study produced results not comparable to the loss function analyses.

When the Committee reran the study using distributive accuracy, its results generally supported adjustment, and this correction was published in the C.A.P.E. addendum.[7] Subsequently, senior Census Bureau statisticians conducted an even more detailed assessment of the effects of heterogeneity on adjustment. They concluded that, even though not in all analysis cases, in general, C.A.P.E.'s analyses showing improvements from adjustment were, in fact, understated.[8]

Status

The Census 2000 A.C.E. design builds upon the C.A.P.E. redesign and incorporates modifications that should reduce or contain heterogeneity and improve accuracy, assuming nonsampling error is reduced or at least contained. First, the Census Bureau has modified the 1992 poststrata to include a mail response variable that should reduce heterogeneity. Second, the larger sample size enables the Census Bureau to improve accuracy at smaller areas. Moreover, adjustment in 1990 and in 2000 has never been intended nor expected to produce substantial improvements in the smallest geographic areas like blocks. Because block estimates of undercount are based on indirect data that is subject to some heterogeneity, some blocks' undercount will be over- and some will be under-stated. Nevertheless, while slightly improving, on average, block-level accuracy, the real benefits of adjustment are in the ever increasing accuracy as blocks are aggregated. Any failure in adjustment to produce substantial improvements for very small areas should not preclude the benefits from adjustment for areas of more reasonable size, such as states, congressional districts, counties, and medium and large cities.

[7]Ibid.

[8]"The 1990 Post Enumeration Survey: Statistical Lessons in Hindsight," Robert E. Fay and John Thompson, *Proceedings of the 1993 Annual Research Conference,* U.S. Bureau of the Census, Washington, DC, pp. 71-91.

Fifth Issue: Can the inconsistency of PES and other estimates be explained?

As part of the July 1991 decision on whether to adjust the 1990 census, a key concern was the inconsistency of PES estimates compared to other estimates, mainly DA. C.A.P.E.'s primary concern was that the PES estimated a higher population than DA and about a million more women. Because the PES was subject to correlation bias, this was an unexpected result. Additionally, face validity checks (informal checks of other estimates of total state population) indicated some areas of concern. Therefore, the Committee requested additional research to investigate the apparent differences. However, after C.A.P.E. revised the undercount estimates (i.e., reduced from about 2.1 to about 1.6 percent) as a result of the computer processing error and the rematch study, most discrepancies were removed. For example,

- The revised PES estimates were now lower than DA (as expected).
- The PES estimates of women remained higher (not expected), but the difference was reduced from about 1 million to about 400,000 and was within sampling error.
- As expected because of correlation bias, the PES estimates for Blacks, especially Black males, were much lower than the DA estimates.
- Face validity checks were also more consistent.

Status

For Census 2000, other estimates such as DA will be used as part of the evaluation of the quality of the A.C.E. Any differences will be included in evaluation reports.

An Analysis of the Consistency of the 1990 Mosbacher Guidelines to U.S. Census Bureau Standards

Sally M. Obenski and Robert E. Fay
U.S. Bureau of the Census
16 May 2000

This paper was prepared in conjunction with the report, *Accuracy and Coverage Evaluation: Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of the Census*, Kenneth Prewitt, U.S. Bureau of the Census, June 2000, available online at http://www.census.gov/Press-Release/www/feasibility.html. The authors' addresses are Sally Obenski, PRED, FOB 2, Rm 1103, U.S. Bureau of the Census, Washington, DC 20233, Sally.M.Obenski@ccmail.census.gov; and Robert Fay, U.S. Bureau of the Census, Washington, DC 20233-9001, Robert.E.Fay.III@ccmail.census.gov.

5/16/00

An Analysis of the Consistency of the 1990 Mosbacher Guidelines to U.S. Census Bureau Standards

Sally M. Obenski and Robert E. Fay

Section I: Introduction

On July 15, 1991, Secretary of Commerce Robert A. Mosbacher announced his decision not to adjust the official 1990 census population counts with the results of 1990 Post Enumeration Survey (PES). His decision document[1] evaluated the PES adjustments against eight previously established guidelines. [2] The first three guidelines specified the following:

- Guideline One required convincing evidence that the adjusted counts be established as superior at national, state, and local levels; otherwise, the unadjusted census counts were to be concluded as the more accurate.
- Guideline Two required that adjusted counts be consistent and complete across all geographic levels, down to the census blocks that constitute the basic units of tabulation.
- Guideline Three required that the adjustment method be specified in advance and that the method be robust, that is, insensitive to small changes in data or assumptions.

The remaining guidelines concerned issues related to the legality of adjustment, the completeness of the adjustment activities and documentation, and the consequences of adjustment for the 1990 and future censuses. The decision document analyzed the evidence in terms of the guidelines and concluded that several of the guidelines, including the first three, supported a decision not to adjust.

Secretary Mosbacher's decision document directed the U.S. Census Bureau to continue analyzing the PES to incorporate, as appropriate, its results into the intercensal population estimates. Subsequently, the Census Bureau improved the PES data and methods and produced a new set of census adjustments in 1992. In early January, 1993, however, Census Bureau Director Barbara Bryant decided not to adjust the population base of the intercensal population estimates with the PES findings, for reasons

[1] 56 Fed. Reg. 33582-33642 (July 22, 1991).

[2] 55 Fed. Reg. 9838-9661 (March 15,1990).

similar to Secretary Mosbacher's Guideline One.[3] Thus, a second precedent was established for considering the first guideline for census adjustments.

Do the Mosbacher guidelines apply as standards for the Census Bureau's plans for Census 2000, including the adjustment of the counts on the basis of the Accuracy and Coverage Evaluation (A.C.E.)? This analysis compares the Mosbacher guidelines and the Census Bureau's practices in designing censuses and surveys. It is the latter set of practices, rather than the Mosbacher guidelines specifically, that have guided the Census Bureau's development of plans for Census 2000. The analysis considers (1) how technical concerns expressed by Secretary Mosbacher would have been altered had the Census Bureau's practices for evaluating census operations been used rather than the guidelines, (2) the consistency of the 1990 guidelines to the Census Bureau's standards, and (3) the status of other technical concerns raised by the 1991 decision document. The focus of the discussion is on the first three guidelines, because of their prominence in the decision document, but the conclusion section will review the more limited relevance of the remaining five guidelines to the Census 2000 planning.

Section II: Secretary Mosbacher's Key Technical Concerns

In the July 1991 decision paper on whether to adjust the census counts, Secretary Mosbacher identified and discussed a number of concerns about the quality of the PES data. The following summarizes his key concerns.

Increased Accuracy at all Geographic Levels: Secretary Mosbacher's principal reason for not adjusting was that the Census Bureau failed to demonstrate increased accuracy at all geographic levels as required by Guideline One. The decision paper stated that there was general expert agreement that the adjusted counts were better at the national level, but there was disagreement about whether the adjustment achieved improvements at lower levels of geography. It further stated that, while the Census Bureau's analyses indicated that more people lived in jurisdictions where the adjusted counts appeared more accurate, one third of the population lived in areas where the unadjusted census counts appeared more accurate. With respect to places under 100,000, the decision paper stated that there was no direct evidence that adjusted counts were more accurate.

Numeric v Distributive Accuracy: One criticism contained in the decision paper was that the Census Bureau was too concerned about *numeric accuracy*, that is, getting the count closer to the true total, rather than *distributive accuracy*, that is, getting the allocation of the population among the states

[3] 58 Fed. Reg. 69-78 (January 4, 1993). For information on the technical issues see "Analysis of C.A.P.E. Findings on the 1990 PES Technical Issues," Sally M. Obenski, June 9, 2000. For information on accuracy, see "Analysis of C.A.P.E. Findings on PES Accuracy at Varying Geographic Levels," Sally M. Obenski and Robert E. Fay, June 9, 2000.

and other geographic units closer to the true proportional distribution. The decision document interpreted *accuracy* required by Guidelines One, Two, and others specifically as distributive accuracy. Secretary Mosbacher concluded that overall distributive accuracy was not improved by the adjusted counts and that there was no convincing evidence the adjusted counts gave a more accurate representation of the distribution of the population across various levels of geography. He stated that the evidence indicated the unadjusted census counts probably yielded a more accurate measure of the distribution of the population.

Number of State Shares Improved: Related to distributive accuracy, a key performance measure that Secretary Mosbacher used under Guideline One to determine which was more accurate–the census or the adjusted numbers--was to count the number of states whose shares improved from adjustment. He stated that the Census Bureau estimated that the proportional share of about 29 states would be made more accurate and about 21 states would be made less accurate by adjustment. The unadjusted census appeared more accurate in 11 out of 23 large metropolitan areas. Secretary Mosbacher noted that as the population units got smaller, including small and medium-sized cities, the adjusted figures became increasingly unreliable. Such measures figured prominently in the decision document's discussion of Guideline One. (The use of number of state shares improved as a performance measure was determined to be methodologically flawed. Section V further discusses this issue.)

Inconsistency with Demographic Analysis: One of the 1990 evaluations compared PES estimates nationally to estimates based on Demographic Analysis (DA). DA constructs an analytic estimate of the population from sets of aggregate data, such as birth, death, immigration, and Medicare statistics, and it consequently offers an alternative method to estimate the net undercoverage of the census. According to the decision paper, the PES would have moved the count of the population in the opposite direction from DA for some demographic groups. Comparisons of coverage measurement surveys for the 1980 census and previous years to DA consistently showed the coverage measurement surveys to underestimate parts of the population. Secretary Mosbacher noted that the PES estimate of the total 1990 population exceeded the estimate from DA. Further, he noted that DA has been the basis of the generalization that males are more subject to being missed than females, but in 1990 about one-half of the people added by the PES were women.

Guideline One had explicitly required the Census Bureau to compare the PES and DA results. The decision document used the claim of inconsistency between them as part of its basis for stating that the guideline had not been met.

Variance of the Estimates: In 1991, a complex statistical "smoothing" model was used to offset unacceptably high variances in the direct PES estimates. Estimates of variance for the smoothed estimates were calculated under a set of assumptions and approximations. But it was not until June 1991 that Census Bureau statisticians identified all the elements of variance associated with the smoothing methodology and provided estimates of how much larger the true variances might be than the variance estimates that had been calculated as planned. Secretary Mosbacher pointed out that, although the analysis based on the available estimates of variance found marginally improved distributional accuracy for adjusted counts, when the analysis was repeated with allowance for potential increases in the variances, the estimated improvements from adjustment diminished and were, hence, unreliable. Indeed, the document described the effect of incorporating plausible estimates of variance as shifting the comparisons toward favoring the accuracy of the census enumeration over the adjustment. For example, he cited Census Bureau estimates that the number of states made less distributively accurate by adjustment would rise from 21 to 28 or 29 states. Additionally, Secretary Mosbacher stated that the statistical tests of whether the accuracy is improved by an adjustment at state and local levels showed mixed results and depended critically on assessments of the amount of variance in the estimates. The decision document used this evidence to support the argument that Guideline One had not been met.

The Synthetic Assumption: The adjusted census counts were produced through synthetic estimation. The 1990 PES design divided the entire country into poststrata--groupings of people defined by age, race, sex, Hispanic origin, tenure, and geographic area. The intention in designing the poststrata was to identify groupings that might differ from each other in the undercount rates. Each individual poststratum was expected to identify a set of people with identical chances of underenumeration. For each of the 1,392 poststrata in the design, the Census Bureau computed a Dual System Estimate (DSE) from the PES data. Then, adjustment factors were calculated as the ratio of the DSE for each poststratum to the census count. The adjustment factors consequently indicated relatively how much a poststratum was over- or undercounted.

The DSE estimate and corresponding adjustment factor for each poststratum were *direct estimates*[4], based on information from the A.C.E. sample for that poststratum; that is, the estimate for each poststratum depended solely on the data for that poststratum[5]. To reduce variance, the adjustment factors were smoothed through the complex statistical model just described. Because the model averaged results across poststrata, the smoothed adjustment factors were *indirect estimates.* But the poststrata did not correspond to geographic areas of primary interest, such as states, counties, and places. Instead, a further stage of indirect estimation, synthetic estimation, was used to produce the adjusted census numbers.

---

[4]Wesley L. Schaible (ed.) (1996) *Indirect Estimators in U.S. Federal Programs*, New York: Springer, particularly Chapter 1, Introduction and Summary.

[5]In fact, missing data adjustments employed data across poststrata, but this is a common situation in survey practice. The distinction between *direct* and *indirect estimators* in Schaible, *op. cit.*, classified estimators for a domain if the estimates depended almost exclusively on data from the domain, even when standard missing data adjustments did use some data across domains.

Synthetic estimation is the logical consequence of what is frequently termed the *synthetic* or *homogeneity assumption.* The assumption is that, at every geographic level (for example, states, counties, blocks) where members of the poststratum are included, they share the poststratum's coverage factor. In effect, the synthetic adjustment assumes that the probability of being missed by the census is constant for each person within a poststratum. The synthetic estimate of the population for any given area was calculated by summing over poststrata the product of the number of persons in each poststratum and the smoothed adjustment factor for the poststratum, and then adding the census counts of persons not in the PES universe.

Secretary Mosbacher questioned under Guideline Two whether the assumptions that underlay the synthetic adjustment were sufficiently valid to produce adjusted counts accurate enough to be useable at a block or district office level. One of the Secretary's special advisors analyzed simulated adjustments at the block and district office level to examine the effects on numeric and distributive accuracy. He reported that distributive accuracy, that is, the shares, suffered much more from the simulated adjustment than numeric accuracy. Consequently, Secretary Mosbacher took the position that a substantial portion, possibly a majority, of district office-size units could be made worse off by adjustment.

Robustness: Another cited concern was the robustness of the smoothing model used to offset high variances of the direct estimates. Guideline Three required the statistical methods used in the PES, such as the smoothing model, to be robust; that is, they were to be insensitive to relatively small changes in assumptions. Secretary Mosbacher noted that the Census Bureau attempted to demonstrate robustness by showing that the set of various population estimates derived from different smoothing methods were broadly similar to the adjusted counts and, as a group, were distinct from the census enumeration. But he contended that the Census Bureau analysis did not consider the similarity in terms of the population distribution of the sets of estimates or whether the variance inherent in those estimates warranted the discarding of the census in favor of one of the particular estimates. The decision paper concluded that these smoothing techniques relied heavily on the particular choice of statistical assumptions that resulted in large changes in adjustment factors when compared to alternatives. The decision paper stated that some of the assumptions may have led to an overstatement of the undercount.

Section III: The Census Bureau Standards

Although not formally documented, practices shared by decennial census planners guide them in deciding on whether a particular operation should be included in the decennial design. These standards for adopting an operation include:

- Operational feasibility--can it be done?
- Cost effectiveness--is there a less expensive way to achieve the same result?

- Valid enumeration--will it result in missed persons being added correctly to the counts (rather than introducing erroneous enumerations)?
- Increased accuracy--will it improve overall coverage and reduce the differential undercount?

While the first three standards are self-explanatory, the last one, increased accuracy, is more complex. In order to identify comparisons between the Census Bureau's implicit concepts of accuracy and the Mosbacher Guidelines, it is helpful to state the Census Bureau's concepts in terms of *numeric* and *distributive accuracy* as used in the July 15, 1991 decision document. (These terms did not appear in the original guidelines, but the decision document's use of them establishes a precedent for interpreting the guidelines in this manner.) As previously noted, *numeric accuracy* pertains to getting the count closer to the true total, and *distributive accuracy,* to getting the allocation of the population among the states and other geographic units closer to the true distribution. In other words, an operation that increases numeric accuracy will move the counts for particular areas or demographic groups closer to their true totals, while an operation that increases distributive accuracy will improve the counts for given areas or demographic groups relative to other areas or demographic groups, in other words improve the estimated proportions.

Both numeric and distributive accuracy affect uses of the census counts, including critical uses such as apportionment, redistricting, and allocation of Federal funds. An additional key use is as a base for the Census Bureau's program of postcensal population estimates throughout each decade; in particular, the monthly Current Population Survey, which provides monthly labor force and unemployment data, and several other surveys depend on postcensal estimates as population controls.

The current apportionment formula can be shown to rely solely on distributive accuracy. In the context of an adjustment potentially affecting reapportionment, a number of authors have concurred with Secretary Mosbacher's emphasis on distributive accuracy.[6] Because the Supreme Court ruling disallowed the use of sampling for apportionment, however, it is appropriate to reassess the relative importance of numeric accuracy and distributive accuracy for the remaining uses. Besides apportionment, the other three broad uses of census data--redistricting, funds allocation, and survey controls--depend on counts that are both numerically and distributively accurate. Both redistricting and funds allocation require distributive accuracy, but funds allocation may also incorporate requirements for numeric accuracy; preparation of survey controls requires numeric accuracy.

---

[6]For example, Barbara Everitt Bryant (1991), "Recommendation to Secretary of Commerce Robert A. Mosbacher on Whether or not to Adjust the 1990 Census," U.S. Census Bureau, June 28, 1991, p. 13; Mary H. Mulry and Bruce D. Spencer (1993), "Accuracy of the 1990 Census and Undercount Adjustments," *Journal of the American Statistical Association*, 88, p. 1083; and Michael L. Cohen, Andrew A. White, and Keith F. Rust (eds.) (1999), *Measuring a Changing Nation,* Panel on Alternative Census Methodologies, National Academy Press, Washington, DC, p. 70.

Although the Census Bureau seeks to ensure that the census is both numerically and distributively accurate, it is numeric accuracy that is primarily considered during the planning process. This is because it is difficult to predict, *a priori*, the effects of a particular census operation on distributive accuracy.

When census operations combine to increase the numeric accuracy of the census, they may approximate the ideal outcome in which a perfectly numerically accurate count assures distributive accuracy as well. Because historical experience suggests that the ideal will not be achieved, the actual situation is one in which decisions about operations to increase numeric accuracy must be made without full knowledge of their effect on distributive accuracy. By itself, an individual operation to increase numeric accuracy may either increase distributive accuracy, leave it about the same, or make it worse. For example, local efforts to increase census participation will have varying levels of success in increasing numeric accuracy. Consequently, the net effect will be that distributive accuracy will be adversely or indeterminately affected. While increases in numeric accuracy are cumulative (assuming those added to the count are not erroneous), distributive accuracy is sensitive to differences in success rates. Secretary Mosbacher provided the following example of how numeric accuracy could be made better but distributive accuracy worse:

> Suppose you observed an enumeration which missed exactly 5 percent of the people in each and every block. Thus, although 5 percent is missed in each and every block, the proportion of the total population in each block is still estimated correctly. Suppose now that you adjusted this enumeration by increasing the counts in half the blocks by 1 percent and increasing the counts in the other half by 5 percent. On average, you would have reduced the undercount of the population by 3 percentage points, thus improving the numeric accuracy of the nationwide total. The numerical accuracy of the absolute level of the count also would have improved for each block. However, the block proportions would now be wrong. Half the blocks would be 2 percent too small and half would be 2 percent too large relative to the average undercount.

Secretary Mosbacher used this example to show that a statistical adjustment could possibly improve numeric accuracy while making distributive accuracy worse. But the illustration equally applies to a coverage improvement program that selectively improves only some areas. Section V further discusses the interpretation of this example.

In practice, carrying out operations to improve numeric accuracy may tend on balance to improve distributive accuracy rather than to harm it. Operations designed to add persons who belong in the census and have been otherwise missed would tend to increase the count in areas where people still remain to be enumerated, while adding negligible amounts to areas that have been virtually completely counted. In effect, the law of diminishing returns bounds the effect of coverage operations in relatively well-counted areas, as long as the operations add only valid persons to the enumeration.

The initial Census 2000 operations (like their predecessors in earlier censuses) are designed to increase numeric accuracy with the intention of achieving as much distributive accuracy as possible within the constraints of traditional enumeration. The Census Bureau expects, however, that a differential undercount of traditionally undercounted groups will reappear in Census 2000 at a level similar to the 1990 census,[7] and that the unadjusted Census 2000 counts will fail to be distributively accurate as a consequence.

Section IV: The 1990 and 2000 Coverage Measurement Surveys

The decision document of July 1991 summarized and reacted to research conducted by the Census Bureau and others. As previously noted, Secretary Mosbacher directed the Census Bureau to continue research on the possible incorporation of PES findings into postcensal population estimates. Accordingly, the Census Bureau continued research guided by this objective, but the research influenced plans for Census 2000 as well. The Committee on Adjustment of Postcensal Estimates (CAPE or C.A.P.E.), formed in 1991, directed research on the use of PES in postcensal estimates and reported results in 1992. During this period the Census Bureau: (1) identified a subset of the PES sample that was reexamined and corrected when necessary; (2) surfaced and corrected errors in computer processing affecting the 1991 estimates; and (3) eliminated the smoothing model employed in the 1991 estimates by replacing the original 1392 poststrata with 357, giving larger sample sizes on average in each poststratum.[8] The Census Bureau then analyzed the evidence on distributive accuracy for the revised estimates. The CAPE only considered issues relevant to the question of postcensal estimation, so that questions on the effect of adjustment on apportionment, redistricting, or block-level accuracy were out of the scope of their study. As also noted earlier, Director Bryant used the same standard as Guideline One, improvement in distributive accuracy at all levels, to decide not to adjust the postcensal estimates in 1992. But she also recognized the special role of numeric accuracy for the population controls for surveys, and allowed the subsequent use of PES results for that purpose.

In planning the Census 2000, the Census Bureau first designed an Integrated Coverage Measurement survey with the objective of adjusting all census data products, including the apportionment counts. The 1999 Supreme Court decision excluding the use of statistical adjustments for apportionment led the Census Bureau to replace Integrated Coverage Measurement with the Accuracy and Coverage

---

[7] The persistence of the differential undercount is illustrated by Table 1 in J. Gregory Robinson, Bashir Ahmed, Prithwis Das Gupta, and Karen A. Woodrow (1993), "Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis," *Journal of the American Statistical Association*, 88, p. 1063.

[8] The history of the 1991 estimates and the changes for the 1992 estimates are described more fully by Howard Hogan (1993), "The 1990 Post-Enumeration Survey: Operations and Results ," *Journal of the American Statistical Association*, 88, 1047-1060.

Evaluation (A.C.E.). The design of the A.C.E. is similar in many respects to the 1990 PES, and its estimation strategy is similar to the 1992 estimates analyzed by the CAPE. The Census Bureau's expectations for the A.C.E. rest in most respects on the CAPE analysis or extrapolation of its results to the larger sample size of the A.C.E. (approximately 300,000 housing units compared to 160,000 in 1990) and other improvements.

To the extent that the 1990 experience is the rationale for the Census Bureau's expectations, then the primary criticisms from the 1991 decision document carry potential implications. The following section reviews the current status of issues raised by Secretary Mosbacher with respect to the A.C.E. and notes where subsequent CAPE research or new features of the A.C.E. have addressed them.

Section V: Consistency of Census 2000 Plans with the Mosbacher Guidelines

Guideline One

Increased Accuracy at all Geographic Levels: As worded, Guideline One imposed a stringent standard on adjustment by requiring evidence that adjustment improved the accuracy of the census counts at all geographic levels. In his decision document, Secretary Mosbacher interpreted the guideline more specifically to require increased distributive accuracy at all geographic levels. The guideline explicitly favored the unadjusted census over the adjustment unless the evidence was convincing at all geographic levels. The guideline specifically excluded adjustment if adjustment could only be shown to be an improvement at some levels, even when there was no evidence of harm at others. Both historically and in the planning of Census 2000, no other census operation has been subject to the standard that positive evidence of improvement in distributive accuracy be established at all geographic levels.

The wording of Guideline One created problems of interpretation for the decision document. Had Guideline One been worded that "The Census shall be considered the official count ..." rather than "The Census shall be considered the most accurate count...," it would have established a clearer distinction between determining the official counts from the 1990 census and reporting of scientific evidence. In some places the decision document concluded that the unadjusted census counts were more accurate, where instead the scientifically correct statement would have been that no significant difference in accuracy between the adjusted and unadjusted counts was detected.

Numeric v Distributive Accuracy: As previously noted, Secretary Mosbacher's assertion that distributive accuracy was of primary importance was in the context of a potential effect of the adjustment on apportionment.

Because the Supreme Court has decided that apportionment will be based on unadjusted counts from Census 2000, the Census Bureau has returned to a balance between numeric and distributive accuracy to decide on the use of adjustment for other purposes.

Number of State Shares Improved: Citing the Census Bureau's results, Secretary Mosbacher used the number of states, 21, whose proportional shares would be made less accurate by adjustment, as a performance measure. The performance measure that Secretary Mosbacher used was inappropriate. Although Census Bureau performed a calculation to determine the number of states improved by the adjustment and reported the results, it later concluded that this measure did not accurately reflect the number of states made better or worse, and this basic conclusion was further elaborated in the statistical literature.[9]

The Census Bureau does not plan the use of this measure in analyzing the adjustment for Census 2000.

Consistency with Demographic Analysis: The decision paper cited inconsistencies between the PES and Demographic Analysis (DA) estimates. However, just as PES estimates have sampling error (variance) associated with them, the DA estimates also have measures of uncertainty[10]. Census Bureau experts tried to quantify the uncertainty through probability references. Because of the uncertainty, the majority of the Census Bureau's Undercount Steering Committee concluded in 1991 that the similarity between PES and DA results was sufficient to support the judgement that the PES--although not perfect--was reflecting real undercounts[11]. Further, the revised estimates produced by the CAPE in 1992 removed some of the apparent inconsistencies between the PES and DA estimates.

In Census 2000, the Census Bureau expects the A.C.E. estimates for adult males, particularly for adult Black males, to underestimate the net undercoverage compared to DA.

Variance of the Estimates: Secretary Mosbacher doubted the reliability of the PES estimates because of uncertainty about the properties of the smoothing model. Although considered necessary to offset the high variances in some of the 1991 design's direct estimates, this procedure was troubling to many senior Census Bureau statisticians. Consequently, the 1990 PES estimation was redesigned in 1992, resulting in an improved design that required no smoothing, and it is the 1992 procedure that provided the basic framework for the A.C.E.

---

[9] Undercount Steering Committee, U.S. Census Bureau (1991), "Addendum to the Report of the Undercount Steering Committee," unpublished report; Mulry and Spencer (1993), op. cit., p. 1085.

[10] Robinson et al. (1993), op. cit.

[11] Undercount Steering Committee (1991), "Technical Assessment of the Accuracy of the Unadjusted versus Adjusted 1990 Census Counts," unpublished report, U.S. Bureau of the Census, June 21, 1991, p. 4.

Smoothing will not be used in Census 2000, and the problem of estimating variances for a complex smoothing will consequently not arise.

Other Technical Concerns: In addition, Secretary Mosbacher questioned the overall quality of the estimates because of analyses conducted by his special advisors that included how correlation bias is measured and accounted for, and the amount of nonsampling error present in the undercount measurement.

Correlation Bias: The DSE recognizes that persons may be missed from either the census or the coverage measurement survey, and it estimates the number of persons missed by both by assuming omissions are statistically independent in the two systems. In effect, the DSE assumes that, within each poststratum, (1) persons have the same probability of inclusion in at least one of the two systems, the census or the coverage survey (*homogeneity*); and (2) inclusion in one system does not affect the probability of inclusion in the other system (*causal independence*). Under these assumptions, the proportion of coverage survey persons also included in the census can be taken as an unbiased estimate of the proportion of all persons included in the census. Dividing the census count (adjusted for estimated erroneous enumerations) by this proportion then provides an approximately unbiased estimate of the true population. To make the homogeneity assumption more likely to hold, this calculation is carried out within poststrata defined by demographic and other characteristics (e.g., age-race-sex, renter versus owner, geographic region).

Because of violations to homogeneity or causal independence, the DSE may misstate, and typically understates, the true population of the poststratum. In the past, including in 1990, DA has shown that such understatement, or *correlation bias,* has appeared for adult Black males and, to a lesser extent, in adult non-Black males.

In the presence of substantial correlation bias for adult males, the estimated undercounts for males and females can be similar for ages where DA shows a differential male undercount. This phenomenon, cited in the decision document as evidence against the concept of correlation bias[12], is actually the outcome of it since, by detecting only approximately the same number of missing males as females in 1990, the DSE was underestimating the number of missing males.

Secretary Mosbacher appeared to use correlation bias as a reason not to adjust. He had two primary concerns about correlation bias. First, because the PES estimated a significant number of persons not included in either system, he concluded that it is not possible to judge whether the adjusted census is distributionally superior to the enumeration simply by accounting for the additional persons estimated by the PES. Second, Secretary Mosbacher cited his special advisor who had concerns about how the correlation bias was allocated to the model that measures the PES' total error.

---

[12]Robert A. Mosbacher (1991), op. cit., p. 33591.

> Since there is no unique way to allocate correlation bias to the model that measures the total error in the PES, this allocation, which critically affects conclusions about the accuracy of the census, was not based on empirical evidence on the distribution of those persons not estimated by either the census or the PES but rather on a formula of convenience.

The Census Bureau has consistently acknowledged that DA has provided evidence that the PES is less than perfect. Correlation bias has appeared, however, for groups with undercount rates consistently higher than average. Thus, the census counts understate the true population for these groups even more than the DSE. That is, the error in the census counts includes both the census undercoverage measured by the DSEs, and correlation bias of the DSEs.

Secretary Mosbacher acknowledged that adjustment goes part of the way toward remedying the undercount of Black males, but then he stated that simply increasing numeric accuracy does not necessarily mean increasing distributive accuracy. In fact, his special advisor stated that the implicit assumption that would be made if the people missed by both the census and the DSE were included in the adjustment would be that they are distributed the same way as the post-adjustment population. The advisor concluded that such an assumption had no empirical foundation.

Because correlation bias remains an issue associated with adjustment, it is useful to elaborate the implications of the advisor's analysis. On the basis of the number of males unmeasured by the 1990 PES, the advisor was in effect arguing that it was impossible to judge improvements to distributive accuracy because of the unknown geographic distribution for the missing males. But the census counts were missing those males as well. The same line of argument leads to the conclusion that it is impossible to determine the effect on distributive accuracy of any other census operation, even after the census is completed, because hypothetical geographic distributions for the missing males leave too much uncertainty about the true distribution.

To restate this observation in terms of the Black undercount specifically, suppose that the Census Bureau had been able to include in the 1990 census cost-effective coverage improvement operations that would have removed the same proportion of the Black undercount as measured by the 1990 PES. Hypothetically, the 1990 census counts would have agreed with the actual 1990 PES findings, and the results of the PES for this hypothetical census would have shown no net undercount. Nonetheless, DA would have still identified the number of missing males and the associated correlation bias. Seemingly, a broad social consensus would have supported such improvements to the census, even if some males remained omitted. The lack of direct evidence on the distribution of missing males would not have been a cogent reason to forgo such hypothetical improvements to the census. Since the missing Black males generally live among other Blacks, it is likely that major improvements to the Black counts increase distributive accuracy.

Secretary Mosbacher's second concern stemmed from remarks of an advisor, who pointed out that there is no unique model of allocating correlation bias for the loss function analysis, and that the Census

Bureau's model was ingenious yet untenable. The advisor provided no alternative upon which to compare, however. The Census Bureau shared Secretary Mosbacher's concern about how correlation bias was allocated to sub-national levels. Consequently, to address the effects of allowances for correlation bias on evaluations of the accuracy of census counts versus PES estimates, subsequent CAPE evaluations were typically run both with and without an allowance for correlation bias. Finally, for the 2000 A.C.E. evaluations, the Census Bureau is examining better measures of and methods for allocating correlation bias sub-nationally for purposes of later analysis.

Level of nonsampling error: As for other nonsampling error (for example, matching error) present in the undercount measurement, the Census Bureau invested significant resources in over 20 evaluations to measure nonsampling error in the PES. The errors reflected both under- and over-estimates of the undercount, so were offsetting. For example, correlation bias resulted in an under-estimate of undercount while false nonmatches resulted in an over-estimate. The results were used to develop a total error model that reflected net nonsampling error so that a fair comparison could be made between the relative accuracy of the PES as compared to the census undercount. To the extent possible, the total error reflected all known sources of error. This comparison provided the basis for the CAPE to conclude that the adjusted data were more accurate nationally, by state, and for areas above 100,000 persons.

Further, the Census Bureau has made a number of changes to the A.C.E. design to decrease nonsampling error from 1990 levels. For example, an automated matching system will be used in 2000 with additional edits built into the software that should reduce matching error from 1990 levels. Likewise, the use of laptop computers for person interviewing will decrease transcription error and missing data. As in 1990, the Census Bureau is planning an extensive set of evaluations to support a total error model to ensure as valid as possible a comparison between adjusted and unadjusted numbers.

Guideline Two

Guideline Two states that the adjusted numbers will be consistent across geographic levels, of sufficient detail for use, and of a sufficiently high quality. Our standards and adjustment methodology are consistent with this guideline. Clearly, the synthetic model consistently applies the correction factors down to the block level, meeting the requirements for consistency and detail.

The Synthetic Assumption: In the decision paper, Secretary Mosbacher concluded that there is no convincing evidence that homogeneity within the poststrata used in adjusting the census counts is a statistically valid assumption. He also concluded that heterogeneity may lead to less accurate counts at local levels. As stated earlier, one of the Secretary's special advisors conducted analyses of simulated adjustments at the block- and district office-level and concluded that distributive accuracy, that is, the shares, suffered much more from the simulated adjustment than numeric accuracy. Consequently, Secretary Mosbacher took the position that a substantial portion, possibly a majority, of district office-size units, could be made worse off by adjustment. The work described in the decision paper was preliminary, however, in three respects. First, the analysis was conducted on only one cluster set out of the 10 in the advisor's sample. Second, under the distributive accuracy simulation, only 7 out of 100 runs of the data showed a majority of districts made worse in relative terms by adjustment. Third, even the advisor himself stated that the results were preliminary.

Prior to the decision paper, the Census Bureau's Undercount Steering Committee analyzed evaluation reports and other independent analyses of heterogeneity. From these results, the majority concluded that empirical evidence suggested that advantages of the adjustment relative to unadjusted counts were not substantially distorted by lack of an explicit evaluation of heterogeneity. Further, the CAPE subsequently addressed the heterogeneity issue in its analyses of the 1992 redesign and concluded that it did not have a deleterious effect on the accuracy of the adjusted counts. Finally, the Census Bureau has adopted an improved poststratification strategy for the A.C.E., taking advantage of the increased sample size to improve homogeneity.

Guideline Three

Prespecification: Secretary Mosbacher concluded that with operations as complex as the census or the PES, complete prespecification was simply impossible. The vast majority of statistical methods were prespecified. However, the Census Bureau consulted with the Secretary's Special Advisory Panel members to ensure their concurrence when the data required deviation from the prespecification. For example, after consultation, Census Bureau statisticians changed a parameter used in the smoothing model. Secretary Mosbacher concurred with the Census Bureau that it must be free to make technically defensible adjustments to the planned process as actual data became available.

For 2000, the Census Bureau will have prespecified most or all aspects of estimation.

Robustness: Although Secretary Mosbacher discussed possible robustness issues about both the imputation model for missing data in the PES and the Census Bureau's poststratification strategy, his main concern was with the lack of robustness associated with the smoothing procedure. The Census Bureau agreed with Secretary Mosbacher that models that are used in producing estimates must be robust. The Census Bureau analysis concluded that the imputation Model was robust and that a lack of robustness from alternative poststratification groupings would not seriously affect the accuracy of the adjusted counts. However, the Census Bureau agreed with Secretary Mosbacher's concern about the smoothing procedure.

Accordingly, the Census Bureau has decided not to use smoothing in creating the estimates for the 2000 A.C.E.

The Other Guidelines

The remaining guidelines are less technical and do not have a clear implication for Census 2000 planning.

- Guideline Four concerned the impact of adjustment on future census efforts. Although Secretary Mosbacher argued that, on balance, adjustment might harm future censuses, his arguments are beyond the technical scope of this document.
- Guideline Five concerned the legality of adjustment, but the decision document took no position on this issue.
- Guideline Six concerned the completeness of the adjustment and the supporting research. For Census 2000, the Census Bureau has committed to producing all estimates required for redistricting by the statutory April 1, 2001 deadline. The Census Bureau is currently planning an evaluation program for Census 2000 but uses the findings from the 1990 evaluations in guiding many of its decisions.
- Guideline Seven concerns the disruptive effect of adjustment. In 1990, the adjustments were available only after unadjusted counts had been released, including counts for redistricting. The disruption stemming from the initial use of unadjusted numbers followed by a potential switch to adjusted ones will be avoided in 2001 by releasing the adjusted counts in time for statutory deadlines.
- Guideline Eight required a full articulation of the basis and implications of the adjustment and adequate documentation for it. The decision document stated that this guideline had been met. The Census Bureau is also addressing this issue in Census 2000.

ASSESSMENT OF ACCURACY OF ADJUSTED VERSUS UNADJUSTED 1990 CENSUS BASE FOR USE IN INTERCENSAL ESTIMATES

REPORT OF THE COMMITTEE ON ADJUSTMENT OF POSTCENSAL ESTIMATES
BUREAU OF THE CENSUS
DEPARTMENT OF COMMERCE
AUGUST 7, 1992

RECOMMENDATION

The Committee on Adjustment of Postcensal Estimates (with an acronym of C.A.P.E. but referred to in this report as the Committee) investigating potential census adjustment for intercensal population estimates concluded that on average, an adjustment to the 1990 base at the national and state levels for use in intercensal estimates would lead to an improvement in the accuracy of the intercensal estimates. (Attachment 1 contains a list of the members of the Committee.) This conclusion was based on a set of extensive research and analyses as well as input from outside consultants. This outside technical advice included a Panel of Experts whose work culminated in a day-long meeting with Census Bureau staff. (Attachment 2 contains a list of the Panel of Experts.) Under the auspices of the Office of Management and Budget (OMB), there also was consultation with other Federal agencies, which are prime users of intercensal estimates.

In coming to its conclusion, the Committee did not vote. Instead, there was an attempt to reach consensus. The conclusion of the Committee was not unanimous, but the large majority of the Committee agreed with the finding. Since there was no vote, this report does not contain a specific listing of minority opinions. Rather, a series of concerns is listed. There was general consensus on several key points.

1. This decision was separate and distinct from the June 1991 decision about whether to adjust the 1990 census for all uses. Making a decision about whether to adjust the full census is quite different from deciding whether to adjust the base that is used in mathematical algorithms to produce estimates of population at several points in the decade between censuses (intercensal estimates).

2. The majority of the Committee concluded that on average, an adjusted state base would be more accurate than an unadjusted state base for use in intercensal estimates, but the Committee recognized there is not necessarily improvement for each and every state base. In fact, the Committee was concerned about a few specific states where the evidence was inconsistent as to whether adjustment was making an improvement. Even so, the Committee felt that overall there was improvement at the state level.

3. States are an important political entity and the first tier in most funding programs. Therefore, the Committee felt that every state or none of the states should be adjusted. Even though some states are smaller than several large cities, the Committee did not recommend adjusting selected cities or counties.

4. For smaller areas (generally, areas of less than 100,000 population), some of the Committee judged that the use of an unadjusted base for the estimates was better than the use of an adjusted base. Other Committee members concluded there was no way to determine whether an adjusted or unadjusted base was more accurate. In the absence of data showing improvement by adjustment, the Committee concluded that the relative distribution of population by substate areas within each state was more

accurate using census counts than the comparable relative distribution using adjusted counts.

5. The Committee was quite concerned about adjusting some, but not all substate areas, especially since there was no way to determine the cutoff of which areas to adjust and there had been no research on the effect of adjustment for a partial set of substate areas.

The Committee's technical assessment was based on a massive amount of data. While there was a re-examination of the information already collected in conjunction with the evaluation of the Post Enumeration Survey (PES), the Committee relied mostly on a large volume of additional research conducted since July 1991. In performing this additional research, the Census Bureau had more time so it could take full advantage of what it had learned from its analysis to date of the 1990 census and the PES. The Census Bureau also had fewer constraints to use prespecified procedures compared to the process in conjunction with the July 1991 decision whether to adjust the 1990 census for which a court order required prespecified procedures. This additional research turned out to be extremely useful, not only for this decision, but for future surveys of all kinds, including those designed for potential adjustment. The Committee wants to acknowledge specifically the massive effort that the professional statistical staff at the Census Bureau put into this research. It was research of such quality that all those involved should be rightly proud. The quality and usefulness of the research also were noted by the set of outside experts that helped review Census Bureau research.

A full description of this research is beyond the scope of this report, but a summary is provided. There are, however, extensive minutes of the Committee meetings, which contain, as attachments, the major results of the additional research. The Committee would like to commend David Whitford and Michael Batutis for preparing these excellent minutes.

In addition to providing useful information, this additional research detected some errors and made some refinements to the levels of estimated undercount originally reported in the spring of 1991. These changes are summarized in the following table and described more fully later in the report.

Population Group	Estimated Undercount			
	June 1991		July 1992	
	Undercount Estimate	Sampling Error	Undercount Estimate	Sampling Error
U.S. Total	2.08%	.18%	1.58%	.19%
Black	4.82	.29	4.43	.51
Asian and Pacific Islander	3.08	.47	2.33	1.35
American Indian, Eskimo, or Aleut	4.77	1.04	4.52	1.22
Hispanic (Can be of any race)	5.24	.42	4.96	.73

This report is a summary of the process that led to the Committee's recommendation. Though the report concentrates on activities that took place late in the decision process, the report also covers several topics that were discussed throughout the year of deliberations by the Committee. Some readers of this report may desire further background on the issue of undercount in a census and the efforts of the Census Bureau to measure and potentially correct (adjust) for any such undercount. There are numerous documents that could be read for background. One good summary document is the notice in the Federal Register concerning the decision of the Secretary of Commerce about whether to adjust the 1990 census (Reference: Federal Register, Volume 56, #140, Part III, pages 33582-33692). The remainder of this report is divided into several sections.

BACKGROUND - UNDERCOUNT	This section contains a description of coverage in the decennial census as well as the methods the Census Bureau uses to measure coverage.
BACKGROUND - ESTIMATES	This section contains a description of why the Census Bureau undertook the task of examining whether to adjust intercensal estimates as well as a very brief description of the estimates program and its use.
RESEARCH	This section summarizes the additional research done since July 1991. This research was the major foundation for the Committee's assessment.
DECISION	This section briefly describes the decision process of the Committee as well as the Executive Staff. These final discussion as well as the year long deliberations of the Committee will be key pieces of input to the Director's decision.
FUTURE	This section contains a few general findings concerning the process of measuring undercount in the future.

BACKGROUND ON UNDERCOUNT

The issue facing the Committee was whether potential error in the PES and adjustment technology was at a sufficiently low level to recommend the inclusion of results from the PES into intercensal estimates. The decennial census is also subject to error, and the PES tries to measure the net coverage error in the census.

This section describes the operations of the 1990 PES to measure census coverage error and how these PES results might have been used for a potential adjustment of the 1990 census. This section is provided solely for background, so the section can be skipped for those already familiar with coverage error in a census as well as the Census Bureau's methods to measure coverage error by the PES and Demographic Analysis.

Since the very first census, there have been problems in accurately counting every person living in the United States. The resulting undercount, or percentage of the population that is not counted by the census, is not a new phenomenon. Beginning with the 1940 census, each decennial census has included an evaluation program to attempt to measure the extent of undercount, or what is often called coverage error. These evaluations showed a steady improvement in net census coverage over four decades, from an estimated undercount of more than 5 percent for the total population in 1940 to an estimated undercount in 1980 of just over 1 percent. They also have shown larger undercount rates for the Black population than the non-Black population and a differential that has stayed about 3-4 percentage points over the period. A difference in estimated undercount for one population subgroup (like Blacks) and another population subgroup (like non-Blacks) is called the differential undercount.

Because of concern about this differential undercount, it was suggested that if the Census Bureau can estimate the number of people missed in a census, why not simply correct the census to account for missed persons and thereby make the census more accurate. This, in simple terms, is what is called "adjustment." But estimating the census undercount with acceptably small error and, in turn, using that knowledge to improve the census counts for all levels of geography are two highly complex and difficult tasks.

The Census Bureau had two major programs to measure coverage in the 1990 census. The first was the PES, which was a sample survey taken after the census. Approximately 165,000 housing units in a sample of 5,290 census blocks or block clusters were interviewed. Block clusters are combinations of small blocks. For the rest of this report, block will be used to mean a block or a block cluster. Persons enumerated during the PES were also referred to as the P-sample. After persons in the housing units in the selected sample blocks were interviewed, their responses were matched to census records in the same set of blocks to determine whether they were counted in the census. This process measured erroneous omissions in the census.

The Census Bureau also measured erroneous inclusions in the census by determining whether any of the persons in the PES sample blocks who were enumerated in the census should not have been counted or should not have been counted at that particular location. An erroneous census enumeration, for example, could have included a child born after April 1, 1990, a person who died before April 1, or a college student away from home who was enumerated at his or her parents' address, instead of being correctly enumerated at the college. Persons in this sample constitute the E-sample.

The data on erroneous inclusions and erroneous omissions were used to produce an estimate of the net undercount or net overcount of the population in the census. This was a very complex process that combined elements of survey design, interviewing, matching, imputation, mathematical modeling and professional judgment.

Second, the Census Bureau used a system called Demographic Analysis (DA) to also measure census coverage. Basically, in DA, an independent estimate of the total population is produced by combining various sources of administrative data. This process included using historical data on births, deaths, and legal immigration; estimates of emigration and undocumented immigration; and Medicare data.

Demographic analysis estimates were used to evaluate the reasonableness of the PES estimates. Only the PES provided estimates of undercount and overcount at a level of detail suitable for use in potential adjustment. For example, demographic analysis estimates were produced only at the national level and for the Black and non-Black populations; the PES process was designed to measure coverage error for more population subgroups (Whites, Blacks, Hispanics, Asians and Pacific Islanders, and American Indians) by detailed levels of geography. Therefore, only the PES data could permit an adjustment.

Each of these programs will be summarized below. For a more detailed discussion of PES see Howard Hogan, "The 1990 Post-Enumeration Survey: An Overview," a paper presented at the American Statistical Association in August 1990; for a more detailed discussion of Demographic Analysis see J. Gregory Robinson, "Plans for Estimating Coverage of the 1990 United States Census: Demographic Analysis," a paper presented to the Southern Demographic Association, in October, 1989.

POST-ENUMERATION SURVEY (PES)

Sample Design

The PES sample was selected in stages. First a random sample of blocks was drawn. Blocks are small polygons of land surrounded by visible features. Most are like the four-sided blocks in a city. Within the selected set of sample blocks, all housing units were listed.

To select the sample of blocks, all blocks in the United States were assigned to one of 101 groups called strata. The strata were defined by geography, city size, racial composition, and percent of housing units that were renter occupied as opposed to owned. A representative sample of blocks was selected from each of the sampling strata. A separate sampling stratum was defined for American Indian Reservations.

Persons living in institutions were excluded from the PES, as were military personnel living in barracks, people living in remote rural Alaska, and persons in emergency shelters and persons who had no formal shelter.

Listing and Interviewing

In February 1990, Census Bureau interviewers who are part of the permanent Census Bureau staff of interviewers visited each of the sample blocks to list all housing units. To preserve independence, none of the temporary enumerators hired to take the 1990 census was used for this listing operation and the listing operation was not conducted out of the temporary census offices. The reason for this was to make sure that temporary people taking the census did not know where a PES sample block was, because if they did, that block might be treated differently during the census.

After the completion of the regular 1990 census interviews, PES interviewers interviewed persons at households in the PES sample blocks. Although this interviewing drew from interviewers who had already worked on the 1990 census, steps were taken to preserve independence, such as not allowing an interviewer to work in a block in the PES that he or she had worked in during the census.

During the PES interview, the interviewers determined who was living in each housing unit, obtained their characteristics, and asked where they lived on April 1, 1990, Census Day. This latter question was necessary in order to determine whether those people who had moved since census day had been counted in the census. The PES interviewing began nearly 3 months after Census Day.

There was a quality assurance program for the interviewing phase to ensure that the interviewers really visited the household and that the people listed were indeed real. If interviewers made up people, they would not match to the census and would inflate the undercount rate.

Matching

The next step was to match the persons enumerated during the PES (the P-sample) to the census. Those persons in the P-sample matched to the census were considered to have been counted in the census; those nonmatched were considered to have been missed.

Matching was carried out in several stages. It involved an initial stage of computer matching followed by clerical matching to attempt to resolve cases that the computer could not match. Many of the persons not matched to the census by computer and clerical matching were assigned for a follow-up interview, if it was determined that additional information might help establish whether a match to the census was appropriate. An additional stage of clerical matching was then conducted using the information from the follow-up interview.

The E-sample, those persons in the PES blocks who were enumerated in the census, was examined to determine if they were correctly enumerated. E-sample persons were matched back into the census to determine if they were enumerated more than once (duplicates). The E-sample persons who were not matched to the

P-sample were potential candidates for erroneous enumerations. Some of these unmatched census persons were also included in the PES follow-up operation described above.

A final matching and reconciliation operation took place at the conclusion of the PES follow-up. An important aspect of this operation was that situations arose where correct match status for persons in the P-sample, or correct enumeration status for persons in the E-sample, could not be determined. This situation occurred because the initial interview was inconclusive or because an incomplete interview was obtained during the follow-up.

Imputation and Dual System Estimates

A final PES computer file was created that reflected the match status for persons in the P-sample and the enumeration status (correct or erroneous) for persons in the E-sample. Computer editing or imputation was performed to correct, insofar as possible, for missing or contradictory data. A critical aspect of imputation involved the estimation of a final match status for those persons whose match status could not otherwise be resolved.

The data in the final PES file were then summarized and incorporated with data from the full census to produce dual system estimates (DSE's) of total population. Dual system refers to the fact that two systems (the census and the PES) are used to make the population estimate. The DSE's were produced separately for each of 1,392 unique subgroupings of the population called post-strata. (See the following section titled Post-strata)

The DSE model to estimate total population conceptualized each person as either in or out of the census cross classified as either in or out of the PES. Essentially it involves determining how many people were (1) in the PES and in the census(matches), (2) in the PES and out of the census(Non-matches), (3) in the census but not in the PES, and (4) in neither the census or PES.

To get an estimate of total population, you could add up the four cells listed above. But, only two of those were directly estimated (cell 1, matches, and cell 2, non-matches). Making some assumptions and using some basic algebra, total population can be estimated without direct estimates for each of the four cells. These operations and the DSE are explained more fully in the Hogan paper cited above.

Post-Strata

The Census Bureau prepared the dual system estimates of the total population for each of 1,392 groupings of people called post-strata. The reason for forming the post-strata was to group persons who had similar chances (probability) of being counted in the census. A person's likelihood of being counted in the census (or in the PES) is called capture probability. The post-strata were defined by census division, geographic subdivisions such as central cities of large metropolitan statistical areas, whether the person was the owner or renter of the housing unit, race, age, and sex. Each person in the PES sample belonged in one of the unique post-strata.

For purposes of illustration, the following are examples of the 1,392 post-strata. One example is a post-stratum which contains Black males, age 20-29, living in rented housing in central cities in the New York primary metropolitan statistical area. A second example is that which contains non-Black non-Hispanic females, age 45-64, living in owned or rented housing in a non-metropolitan place of 10,000 or more population in the Mountain Division. A third example is that which contains Asian males, age 45-64, living in owned or rented housing in metropolitan statistical areas but not in a central city in the Pacific Division. A fourth example is that which contains non-Black Hispanic females, age 30-44, living in owned or rented housing in central cities in the Los Angeles-Long Beach primary metropolitan statistical area or other central cities in metropolitan statistical areas in the Pacific region. As can be seen from these examples, the 1,392 post-strata are very specific.

Adjustment Factors

The next step in the process was to compare the estimated total population for each post-stratum (the dual system estimate or DSE) to the census count to determine a "raw" adjustment factor. For example, if the DSE for a particular post-stratum was 1,050,000 and the census count was 1,000,000, then the adjustment factor was 1.05, reflecting about a 5 percent estimated net undercount. Though most adjustment factors are larger than one, indicating an estimated undercount, an adjustment factor may be less than one, which would have the effect of lowering the census count for the post-stratum if an adjustment is applied. This situation results when there is evidence of an overcount in the post-stratum.

"Smoothing" the Adjustment Factors

The next step was "smoothing" these "raw" adjustment factors to reduce sampling variance and to produce final adjustment factors. Because the PES was a sample, it was subject to sampling error. Sampling error is the error associated with taking some of the population (a sample) rather than all of the population (a census). The process of smoothing the "raw" adjustment factors to create final adjustment factors was a step to minimize the effect of sampling error. Basically, smoothing is a regression prediction model. A multi-variate regression using items correlated with undercount predicts the undercount for each of the 1,392 post-strata. Then, the final adjustment factor is an average of the "raw" adjustment factor and the predicted adjustment factor. For a post-stratum with low estimated sampling error, there was heavy weight on the "raw" adjustment factor in the averaging, and vice versa. The smoothing technique was based on certain assumptions and would add an additional component of error called model error. The Census Bureau hoped that the reduction in sampling error from smoothing would offset any additional errors from the smoothing model chosen. If the Census Bureau had not used smoothing, the final adjustment factors for some of the post-strata would have been based on estimates of undercount that were subject to very large sampling error.

Small Area Estimation

The Census Bureau used the final adjustment factors to produce adjusted counts for every block in the Nation. The PES can only produce "direct" estimates of the total population for relatively large geographic areas (*i.e.*, the 1,392 post-strata). If there had been a decision to adjust, however, the adjustment would have been applied to each of the Nation's approximately 5 million populated blocks. The Census Bureau developed a model that took the adjustment factors produced for each of the 1,392 post-strata areas and used them to estimate adjustment counts for each block. Since each of the post-strata contain many blocks parts, the Census Bureau based its model on a critical assumption that coverage error is similar for all blocks parts within a post-stratum. (A block part is simply that part of the block that falls within the definition of a post-stratum. For example, females within a block would be part of a block and in one set of post-strata while males within a block would be in different set of post-strata.) This assumption of all block parts within a post-stratum being alike (homogenous) with regard to the chance of being counted is analogous to the homogeneity assumption for persons.

Finally, the Census Bureau produced a set of census tabulations with adjusted counts. It did this by adding or subtracting "adjustment" persons with detailed characteristics. The number of people added or subtracted was determined by final adjustment factor for the post-stratum that the block part was in. If someone had to be added, the information from someone else in the block part who was counted in the census was duplicated. If someone had to be subtracted, the information for someone in the block part who was counted in the census was deleted.

Evaluations

The PES and adjustment process are based on many assumptions and have the potential for error. To evaluate the assumptions and potential error, the Census Bureau conducted numerous studies called P-studies because they referred to the PES. The studies were associated with the following general areas.

- Missing data on the PES questionnaire
- Misreporting of census day address on the PES questionnaire
- Fabrication of data in the PES by interviewers
- Errors in matching
- Errors in determining erroneous enumerations
- Balancing omissions with erroneous enumerations
- Correlation Bias (the tendency of the DSE to underestimate total population because some people are missed in both the PES and the Census)
- The homogeneity assumption

The results of these evaluations are essential to determining whether adjusted or unadjusted census counts are more accurate.

DEMOGRAPHIC ANALYSIS

The Census Bureau's other coverage measurement program was demographic analysis (DA). DA uses historical data on births, deaths, and legal immigration; estimates of emigration and undocumented immigration; and medicare data to develop an independent estimate of the population. The DA estimate of population is compared with the census count to yield another measure of net census coverage. DA can be only used to make reliable estimates at the national level. The DA coverage estimates were compared to the post-enumeration survey coverage estimates to assess the overall consistency of the two sets of estimates at the national level.

Birth and death records are available for the entire United States from 1933 on, but are not complete for years before 1933. Therefore, the Census Bureau had to find other ways to estimate the number of people who were born or died prior to 1933. In estimating births for each year, The Census Bureau added to the number of registered births an estimate of under-registration. Under-registration was estimated based on tests conducted in 1940, 1950, and 1964-1968. If the estimates of under-registration are off, they could have a significant effect on undercount estimates because birth data are by far the largest component in estimating the population through demographic analysis. Since national birth and death records are not available before 1933, the Census Bureau had to find other ways to estimate the size of the population 55 and older. For the population 65 and older, medicare estimates are used. For the population 55 to 64, estimates are made from revisions to earlier estimates.

The United States does not keep emigration records. Therefore, an estimate had to be made of persons who have left the country. While the United States does have good records of legal immigration, there is no accurate estimate of illegal immigration. The Immigration and Naturalization Service now collects different information than it did prior to 1980. That change further complicated the effort to estimate legal immigration. Also recent legislative reform allowing amnesty also complicated the issue since the Census Bureau did not know whether all of those obtaining amnesty actually reside in the United States. The Bureau used professional judgment to estimate the components of illegal immigration.

It is important to emphasize that results of demographic analysis are not exact but are estimates. To a large extent, they were based on assumptions and best professional judgment. As in the PES, the Bureau tried to estimate potential error in the data produced by demographic analysis in a series of studies call D-studies. Based on these studies, the Census Bureau developed a range of error around the demographic analysis estimates.

UNDERCOUNT STEERING COMMITTEE

To address the evaluation of the coverage in the census and the methods used to evaluate that coverage (the PES and DA), the Census Bureau formed the Undercount Steering Committee (USC). Their work was an important part of the July 1991 decision whether to adjust the full 1990 census for all uses. The work of the USC was also the major basis for the work done by CAPE. For a detailed description of the findings of USC, see Technical Assessment of the Accuracy of Unadjusted versus Adjusted 1990 Census Counts: Report of the Undercount Steering Committee, June 21, 1991.

BACKGROUND ON INTERCENSAL ESTIMATES

When the Secretary of Commerce announced his decision on July 15, 1991, not to adjust the 1990 census, he indicated his concern about the differential undercount. Because of that concern, he instructed the Census Bureau to continue its research into the area of potential adjustment. If the Census Bureau was able to resolve the technical problems associated with adjustment that were identified in the spring of 1991, then the Secretary asked the Census Bureau to consider incorporating results from the PES into the intercensal estimates program.

Basically, intercensal estimates are made by updating the most recent census base with estimates of population change (births, deaths, and net migration). Of course, the actual procedure is much more complicated and sophisticated. The Census Bureau makes estimates at the national, state, and county level every year and at the incorporated place (city) level every other year. These estimates have a variety of uses. Most notably, the estimates are used in funding allocations, as sample survey controls, and as denominators for many important statistics.

About one-third of the Federal funding programs use intercensal estimates of population as part of their funding formula, rather than using the 1990 census count for ten years. There may be items other than total population in the formula as well. The General Accounting Office has estimated that about 10 billion federal dollars a year are allocated based on funding formulas that use intercensal estimates[1]. States have within state fund-allocation programs as well. Many states use intercensal estimates to allocate within-state funding dollars.

Many sample surveys use national, and to some extent state, intercensal estimates as controls. The most notable is the monthly unemployment survey (the Current Population Survey, or CPS). Sample surveys generally have poorer coverage than a census; therefore, in order to improve the accuracy of estimates from a sample survey, the sample survey estimates are often controlled to an independent total (in this case, the intercensal estimate).

Many Federal agencies produce statistics per 1,000 persons (or some other base). Examples are crime statistics, incidence of certain health conditions, etc. The numerator of these statistics can be obtained at various points in time throughout the decade. In the absence of any updated information, calculating these kinds of statistics on a static 1990 denominator would be misleading; therefore, these Federal agencies use intercensal estimates of population as the denominator.

In order to be responsive to the Secretary's request on intercensal estimates, the Census Bureau formed the Committee to address the technical issues related to a potential adjustment of the base for intercensal estimates. The Committee was made up of many people who also served on the Undercount Steering Committee for the July 1991 decision. However, the Committee also

---

[1]Federal Formula Programs - Outdated Population Data Used to Allocate Most Funds (GAO/HRD-90-145, September 1991).

included some new members, including some Census Bureau staff very familiar with intercensal estimates. Though the Committee focused on the technical issues surrounding a potential adjustment, early in the Committee's deliberations, the Committee also had to make some key decisions related to the unique nature of the intercensal estimates program. The Committee decided that:

1. For the purpose of survey controls, there would be only one decision point in the decade about whether to adjust intercensal estimates.

2. If there was a decision to adjust, there would have to be a mechanism to make the intercensal estimates additive from the smallest area to the national total.

3. There would not be adjustment for some uses of intercensal estimates, but no adjustment for other uses of the estimates.

4. If there were a decision to adjust, the amount of the adjustment would be calculated on the base population. This adjustment plus an estimate of population change for the time period since the census would be added to the unadjusted base.

After every census, there is a change in the base used to calculate the intercensal estimates. Apart from the question of adjustment, there would be a change from a 1980 census base to a 1990 census base. For the use of estimates as survey control totals, that changeover date was postponed from January 1992 to January 1993. Therefore, 1992 estimates released in January 1993 would reflect the 1990 base. The postponement was made so that the decision on whether to adjust the base for intercensal estimates could be made at the same time. If there is a decision to adjust, then the change to a 1990 base and the change to a 1990 adjusted base would be simultaneous. If the decision is not to adjust, then there will be a change to the 1990 unadjusted base. In that case, even if evidence later in the decade would lead one to support adjustment, the base would not be changed from 1990 unadjusted to 1990 adjusted at a later point in the decade for the purpose of survey controls. Any change in base presents a discontinuity in uses based on intercensal estimates. Federal agency users of intercensal estimates for survey controls were quite clear that they strongly preferred only one such discontinuity during the decade.

On a technical basis, it is conceivable to be able to support adjustment at one level (say states), but not at lower levels. In such a case, state estimates would add to the national estimate, but substate estimates would not add to state estimates. There was agreement from users and from the staff making the estimates that failure to have additivity was not only undesirable, but close to unacceptable. Also, on a technical basis, it is conceivable to be able to support adjustment for one purpose (for example, national survey controls), but not for another (for example, subnational fund allocation). The Committee found this situation undesirable. Finally, it is possible for the Census Bureau to decide not to adjust the base of estimates but for some Federal agencies to do their own adjustment. This topic was discussed among Federal agencies at a meeting at the OMB. There was general agreement that it would be unacceptable to have variable sets of intercensal estimates used differently by different Federal agencies.

Estimates start with a base population and add estimated population change (births, deaths, and net migration). If estimates are adjusted, an additional term would be added that represents the net adjustment level for each area. This net adjustment level is the difference between the adjusted base population and the unadjusted base population. In the estimation process, the sum of this net adjustment and the estimated population change would be added to the unadjusted population base. Under this procedure, the net adjustment would remain constant throughout the decade.

FURTHER RESEARCH
THE BASIS FOR THE ASSESSMENT

When discussing the issue of whether to adjust the 1990 census, almost all experts agreed that with more time, there would be refinements and changes to the estimated undercount. Most experts, however, assumed these changes would be relatively small. Since the July 1991 decision, the Census Bureau had the time and at the direction of the Secretary of Commerce, continued to examine the estimated undercount. As expected, the Census Bureau has made some refinements and changes. During this analysis, the Census Bureau discovered a significant computer processing error in the system used to determine the undercount estimates that were under consideration in spring 1991. As a result of an error in computer processing, the estimated national undercount rate of 2.1% was overstated by 0.4%. After correcting the computer error, the national level undercount was estimated to be about 1.7%. After making other refinements and corrections, the national undercount is now estimated to be about 1.6%. Attachment 3 shows revised undercount estimates by selected age-sex-race categories. Attachment 4 shows revised undercount estimates by state. Attachment 11 shows revised undercount estimates for cities of 100,000 or more population. Attachment 12 shows revised undercount estimates for counties of 100,000 or more population.

Since PES undercount estimates were based on a sample survey, they are subject to error. There is sampling error to reflect the fact that the information came from some and not all of the population. The estimates are also subject to biases. For example, errors in matching, erroneous responses from respondents, etc. can bias the undercount estimate. Just as for the estimate of undercount, the Census Bureau also refined its estimates of bias. The level of total bias, excluding correlation bias[2], on the revised estimate of undercount is negative 0.73 (-0.73%). Therefore, about 45% (0.73/1.58) of the revised estimated undercount is actually measured bias and not measured undercount. In 7 of the 10 evaluation strata[3], 50% or more of the estimated undercount is bias. When correlation bias is included, these percentages go down. With correlation bias, the revised estimate of total bias is negative 0.35 percent (-0.35%). Including correlation bias, about 22% of the revised estimate of undercount is actually bias and not measured undercount. In general, the Committee was concerned that the estimate of correlation bias could be an underestimate, which meant the total bias estimate of negative 0.35% was an overstatement. There was limited time and methodology to investigate this concern further. The Committee did not feel lack of more information on this concern had an appreciable effect on their overall conclusion.

---

[2]Correlation bias is a term that reflects the fact that the DSE of total population based on the PES is an underestimate for the model used by the Census Bureau. The DSE is downwardly biased because of correlation bias which occurs, for example, because there are people missed in both the census and the PES. Correlation bias is described more fully below in the section entitled Third Issue-Part B, p 21.

[3]See Attachment 6 for a description of evaluation post-strata.

When the Committee began discussing the issue of whether to adjust the base for intercensal estimates, it started by reviewing the technical concerns raised about whether to adjust the 1990 census. This analysis produced a list of concerns, which the Committee summarized into five key areas.

1. Could the problems in the smoothing model, including lack of robustness, be resolved?

2. Could the estimated biases in the PES estimate of undercount be removed?

3. Were all components of the bias adequately reflected in the total error model, and was total error being accurately handled in loss function analysis?

4. Could we learn more about whether or not our homogeneity assumption held sufficiently to support adjustment?

5. Could we resolve the inconsistencies between the PES and other estimates of undercount, primarily Demographic Analysis?

There were other issues raised. While it would have been helpful to research these other questions as well, the Committee felt comfortable in confining its research efforts to the five key questions. The Committee felt they could make a reasoned choice about whether to adjust the base for intercensal estimates if they got appropriate information on these five issues.

FIRST ISSUE: COULD PROBLEMS IN THE SMOOTHING MODEL BE RESOLVED?

Summary: The Committee was very comfortable with the new post-stratification scheme which reduced sampling variance enough to avoid the use of smoothing. However, because of the limitations of artificial population analysis[4], there was still some concern with the finding that there was no loss in homogeneity[5] in a smaller post-stratum design that had only about 25% as many post-strata. (See fourth issue.)

For the July 1991 decision on whether to adjust the 1990 census, the sample of about 400,000 people was post-stratified into 1,392 groups. A person could be in one and only one of the 1,392 post-stratum groupihgs. Some of

[4]Artificial Population Analysis refers to the study to examine if the persons within each of the 357 post-strata were alike (homogeneous) with regard to their probability of being counted in the census. Artificial Population Analysis is described below in the section entitled Forth Issue, p 25.

[5]To make estimates from the PES, each sample person is assigned to one and only one post-stratum. A necessary assumption is that every person within a post-stratum has approximately the same chance of being counted in the census or the PES. This assumption is called the homogeneity assumption.

those post-stratum groupings were quite small so the estimate of undercount was subject to very high sampling variance. In order to reduce this sampling error, the Census Bureau used a technique called smoothing. Smoothing was a regression prediction model. Based on items correlated with undercount, the undercount for each of the 1,392 post-strata was predicted using the regression model. Then, the final undercount was an average of the predicted undercount and the directly observed undercount.

The smoothing process was successful at reducing the sampling variance. However, there were several issues raised about the entire smoothing process. It would have taken a large, intense, and uncertain research program to have answered all of these concerns. Therefore, the Committee chose a different approach. The Committee agreed to reduce the number of post-strata. By doing so, each new post-stratum would have more sample size than under the 1,392 system, and presumably, enough sample size so that the estimates would be stable (meaning the estimates would not have very large sampling variance); therefore, no smoothing would be required. It was expected that there would be some loss of homogeneity by going to a smaller post-stratum design, since with fewer strata, each stratum now had more people. Therefore, one could expect that it was less likely that everyone within these larger strata had the same capture probability as in smaller strata. The Committee assumed that the loss in homogeneity would be smaller than the problems and potential error from smoothing. As it turned out, the Committee's assumption seemed to be correct.

Based on measures of census performance and general patterns of undercount, a new set of 357 strata were designed. The 357 strata were not a simple regrouping of the 1,392 strata. The 357 strata design included 51 main strata defined by geography, owner-renter, and race/Hispanic cross classified by 7 age groupings cross classified by male-female. Attachment 5 contains a description of the 357 post-stratum design. This 357 design turned out to be a very effective stratification, primarily because we were able to examine additional data before defining the strata. Perhaps the most important piece of information for this examination was the strong relationship of living in owner or renter housing units to undercount. Hence, owner-renter status is very prominent in the 357 design.

We prepared revised PES estimates of undercount based on the 357 design and analyzed sampling variance by post-stratum. The intent was to verify the assumption that the sampling variances under the smaller (357) design would be relatively stable. At the state level, the variances were at an acceptable level[6]. Attachment 10 contains revised estimates of undercount or overcount for the 51 main post-strata that were part of the 357 post-stratum design.

The Committee was also concerned with the potential loss of homogeneity with the smaller post-stratum design. Using artificial population analysis, the Committee examined the homogeneity of the 1,392 design compared to the 357 design. Artificial population analysis is described below in the section called Fourth Issue. Based on the artificial

---

[6]C.A.P.E. minutes 4-6-92, Attachment 3.

population analysis assuming no bias in the PES, the Committee found the homogeneity for the 1,392 design and the 357 design to be about the same[7]. This result at first seemed counter-intuitive since one would have expected some reduction in homogeneity. However, the result may be explained by the fact that the 357 design is much more effective than the 1,392 design (probably true since the 357 design was based on a careful review of auxiliary data), by limitations of the artificial population analysis, or by a combination of both those factors.

In summary, the Committee was very comfortable with the new stratification. In general, for state-level estimates, the Committee felt satisfied with the 357 design without smoothing versus the 1,392 design including smoothing. However, because of the limitations of artificial population analysis, there was still some concern with the finding of no loss in homogeneity by going to a smaller post-stratum design that had only about 25% as many post-strata.

SECOND ISSUE: CAN ESTIMATED BIASES BE REMOVED FROM PES ESTIMATES?

Summary: One of the first steps in further analysis of the PES was to re-examine the 104 blocks which had the greatest effect on the undercount. Many of the blocks had such a significant effect, they could be considered outliers. As a result of the examination of 104 blocks[8], corrections to the Post Enumeration Survey (PES) undercount estimates and bias removal were conducted. The net result was to reduce the estimated national net undercount by 0.1%. During that analysis, the Census Bureau also found and corrected a computer error that had incorrectly overstated the 2.1% undercount reported in July 1991 by .4%. The July 1991 estimate of undercount was reduced by 0.4% because of the computer error and an additional 0.1% because of modifications and bias removal resulting in a revised July 1992 national PES estimate of undercount of about 1.6%. The Committee obviously was satisfied that the decision to do a review of 104 blocks led to the discovery of the computer processing error. The Committee was also confident that outlier blocks had been more appropriately handled. As for bias removal, the Committee had mixed feelings. They were pleased that the review of only 104 blocks had removed a relatively large amount of bias. But, a significant amount still remained. The Committee could find no reliable or expedient method to remove the balance of the bias from the PES estimates.

The PES estimates of undercount are subject to biases. The Census Bureau had many evaluation programs to try to measure the level of these biases. At the U.S. level for total population, the estimated bias was negative 0.73% (or negative 0.35% if correlation bias is included) on an estimated

---

[7]C.A.P.E. minutes 4-6-92 Attachment 5 and C.A.P.E. minutes 3-9-92 Attachment 1.

[8]Small blocks were often combined to form block clusters. This report uses blocks to refer to blocks and block clusters.

undercount of about 1.6%. If it was possible, it would be desirable to remove these biases before any potential adjustment since the PES estimate of undercount including the bias is an overstatement of the undercount the PES actually measured. At the U.S. level for total population, the bias could be removed. The Committee discussed the possibility of removing the bias at sub-national levels. The only alternative was a modeling approach. Considering the very small samples used to estimate the biases and the difficulties of modeling, the Committee was very reluctant to try to remove the bias by modeling. The Committee was concerned that more error would be introduced than the level of error we were trying to remove. A further complication was the concern that our estimate of correlation bias was conservative (see page 15).

As a partial solution to bias removal, the Committee recommended an examination of the blocks that had the potential to contribute the most to the PES estimate of undercount. If the bias could be removed from these blocks, the PES estimates would be improved. Of course, the results from this set of blocks could not be generalized to other blocks, so any solution would only be a partial removal of the bias. 104 blocks were included in the study. The study is referred to by various names since additional components to the study were added over time. This study was originally called OCR (Outlier Cluster Review) because of the intent to review the blocks that had outliers. When the study was expanded to a second purpose (removal of bias), the study was called Selective Cluster Review (SCR).

During the SCR, several types of problems were examined. The treatment of outliers was reexamined and corrected as necessary. Some blocks had unusual results and had very big effects on the estimated undercount, effects far larger than one block should be expected to have. These are called outliers. They are similar to unusual marks by judges in athletic competitions. For the July 1991 estimates of undercount, there was a method to defuse the effect of these outliers. Now, with more time, we were able to reexamine these outliers and to use better methods (when applicable) to dampen their effect.

In addition, during SCR, we looked for errors. An example is failure to search in the proper block. Searching for matching should have been done in the PES sample block as well as the ring of blocks surrounding the sample block. Generally, this was done. Sometimes errors were made and the matchers failed to look into the entire ring. Mistakes like these were corrected.

Matching, even in the proper set of blocks, is error prone. Errors in matching can lead to a bias in the PES estimates. During SCR, expert matchers tried to remove all matching error and therefore any bias in the PES estimate due to matching.

As a result of all aspects of SCR, the estimated national undercount was reduced by one-tenth of one percent (0.1%)[9]. The bias reduction only applied to the 104 blocks and could not be generalized to other blocks. The 104 blocks represent about 2% of the total sample while the 0.1% reduction on an estimated 0.7% total bias represents about a 14% reduction. Even though total bias could not be removed, these numbers show that the effort of redoing these 104 blocks was well worth it. The results of the SCR were also subtracted as appropriate from the total bias so that the resulting total bias only represents residual error for residual blocks (the total minus these 104 blocks).

During the SCR, Census Bureau staff discovered a computer processing error that affected the estimates of undercount released in July 1991. Codes that were attached to cases in clerical processing were incorrectly fed into the computer processing. Errors went in both directions (increasing and decreasing the estimated undercount), but the net result of the error was to reduce the estimated national undercount of 2.1% by 0.4%.

THIRD ISSUE: IS THE TOTAL ERROR MODEL COMPLETE?

Summary: With regard to total error, the Committee was completely satisfied that all components of bias were represented. The Committee was concerned about the accuracy of some of the estimates of bias and the high variance for some estimates of bias. The general conclusion was to use caution in evaluating the results of loss function analysis since the target numbers in that analysis were so dependent on the levels of estimated bias. The Committee felt that correlation bias should be a component of total error. However, there was concern about our method of estimating it and very serious concern about the method of allocating it to states, cities, etc. Since there did not appear to be methods or time to analyze this allocation issue further, the Committee requested that loss function analysis be done with and without correlation bias. There was a choice of various loss functions. Primarily, the Committee concentrated on loss functions that examined proportionate population shares and not population counts. In addition, in general, the Committee considered loss functions based on squared error not absolute error. Using hypothesis tests with 10% significance, loss function analysis excluding correlation bias does not support adjustment. Using hypothesis tests with 10% significance and including correlation bias, all but one of the loss function analyses favors adjustment at the state level when examining aggregate loss. The Committee tended to accept these findings keeping in mind the numerous caveats. As a result of some comments from the Panel of Experts, the Committee was concerned about whether the significance level they used for the hypothesis tests was appropriate.

---

[9]"Post Census Rematching for the Outlier Cluster Review," Howard Hogan, undated; C.A.P.E. minutes 6-11-92 Attachment 1,2; C.A.P.E. minutes 4-20-92 Attachment 2.

THIRD ISSUE-PART A: TOTAL ERROR

The third major concern was whether the total error model contained all components of error and whether the components of error were adequately measured. In terms of whether all components of error were considered, two new components were added-- error due to cases done very late in the regular census (called late-late returns) and treatment of out-of-scope cases. The Committee felt completely confident that all components of error had been listed and considered.

The Committee could come to no agreement about the adequacy of the level of error measured for each of these components. There were concerns that matching error was determined by a dependent study and not an independent study. There were concerns that evaluation interviews used to determine the quality of the PES were conducted in February 1991. ten months after the census. There was concern that the estimate of only 13 fabrications in a sample of 150,000 seemed low compared to reasonable expectations. The Committee strongly agreed that the evaluation sample sizes were too small. The sampling error on several of the estimates of bias was extremely high.

In summary, with regard to total error, the Committee was satisfied that all components of error were represented. The Committee was concerned about the accuracy and variance of the estimates of bias, but there was really nothing that could be done. The general conclusion was to use caution in evaluating the results of loss function analysis since the target numbers in that analysis were so dependent on the levels of estimated bias. Attachment 6 contains estimates of the bias.

THIRD ISSUE-PART B: CORRELATION BIAS

The Committee spent a good deal of time discussing one aspect of total bias--correlation bias[10]. The Dual System Estimate (DSE) of total population produced by comparing the PES and the census is a biased estimate. It is biased because of matching error, etc. These components of bias are described immediately above.

The DSE can also be biased by correlation bias which has multiple components. The first is that the DSE assumes that a person's participation in the PES is not affected by his or her participation in the census (the causal independence assumption). Failure of this assumption can cause a bias. Generally lack of independence is not considered to be a big problem since the PES is conducted almost 4 months after the census and because of other controls introduced into the PES system.

The second component of correlation bias occurs because of variable capture probabilities within a post-stratum. The DSE does not require that the census and the PES have the same probability of counting people (called capture probability). But, the DSE does assume that within a post-stratum,

---

[10]Sometimes, model bias is used synonymously with correlation bias. In this report, correlation bias will be used.

everyone in the PES (or everyone in the census) has approximately the same capture probability. So, for example, a white male renter age 30-49 in rural areas of Louisiana is assumed to be just as likely to be counted as a white male renter age 30-49 in rural Mississippi, etc. Generally, if people within a post-stratum have differing capture probabilities, then the DSE is downwardly biased. That means the DSE underestimates the total population and in most cases would underestimate the undercount.

As a special case of variable capture probabilities, assume within a post-stratum there is a set of people with zero probability of being captured. These are often called the impossible to count or people missed in both the census and the PES. They are another component of correlation bias.

There are no direct estimates of either of these components of correlation bias, but an estimate for the total of both combined is obtained by comparing PES estimates to Demographic Analysis (DA) estimates. To estimate the level of correlation bias, the assumption is that sex ratios as determined by DA are accurate. Then, since in general the DSE estimates of males are lower than the DA estimates of males, there is a calculation of how many males would have to be added to the DSE to make the PES sex ratio equal to the DA sex ratio. These added males are an estimate of the level of correlation bias in the PES.

Actually, after estimating the extent of correlation bias, it is not added to the DSE of total population (just as other estimates of bias are not subtracted). Rather, the estimate of correlation bias is added to the total error model and is used to determine target numbers for loss function analysis.

The Committee was concerned about the combination of the two components of correlation bias, but there did not appear to be any alternative. The Panel of Experts expressed the same sentiment. They agreed that they were uncomfortable with the combination, but there does not seem to be an easy alternative. The Committee also was concerned that the PES measures more females than DA so that this method of estimating correlation bias should have had the effect of estimating a true population (for loss function analysis target numbers) that was bigger than total population in DA. However, the sum of the target populations did not equal the sum of the PES estimate and the level of correlation bias that was estimated to be added, as it should have. There was no time to examine these concerns further. Finally, there was concern that the method used for comparing the DSE with bias to DA understated the estimate of people missed due to correlation bias.

Mostly, however, the Committee was concerned with the method of allocating the correlation bias. Basically, the estimated missing people due to all types of correlation bias (all males) are allocated back to each post stratum proportional to the estimate of the number of males in the fourth cell of the DSE for the post-stratum. Further modeling is used to allocate the total error down to sub post-stratum levels.

The fourth cell in the DSE is an estimate of the number of people missed in both the PES and the census, but it is a biased estimate because of correlation bias. It is not directly estimated, but an estimate can be obtained by subtraction. Some of the numbers used in the subtraction are sample estimates, therefore, they are subject to sampling variability. The fourth cell is expected to be the product of the true population times one minus the capture probability of the PES times one minus the capture probability for the census. In theory, this number cannot be negative. But, in practice, due to sample variability, matching error, etc., it can be estimated to be negative. When the estimate in the fourth cell is negative, no amount of the estimated people missed due to correlation bias is allocated to that post-stratum.

Both the Committee and the Panel of Experts were very concerned about the negative values in the fourth cell. The Panel of Experts suggested some methods to change the DSE process to avoid negative values. There was also considerable concern about using the fourth cell as the basis for allocation of the estimate of people missed due to correlation bias. In fact, other methods of allocation had been tried by the Census Bureau.

In summary, the Committee felt that correlation bias should be a component of total error. However, there was concern about our method of estimating it and very serious concern about the method of allocating it. Therefore, the Committee requested that loss function analysis be done with and without correlation bias. Each Committee member would then have to make some judgements about how to analyze the results.

THIRD ISSUE-PART C: LOSS FUNCTION ANALYSIS

Estimates of bias in the PES estimates of undercount are useful for interpreting the accuracy of the PES estimates. But, estimates of bias were also a key component in a summary analysis called loss function analysis. If truth were known, the census count and the adjusted base count could be compared to truth and an appropriate choice could be made. That of course is impossible. To approximate that comparison, the Census Bureau performed loss function analysis.

As a first step in loss function analysis, the true population is estimated. This estimate is called the target population. It is estimated by taking the PES estimate of population and modifying that estimate based on the estimates of error in the PES (the components of bias from the total error model). These estimates of bias are also subject to error, so you can't simply subtract bias from the PES estimate and assume that is the true population. A further complication is that estimates of bias are only available for 10 evaluation post-strata and target numbers are needed for every state, every county, every place, etc. A modeling system is used to allocate the bias from the 10 evaluation post-strata to sub-levels of geography. Once target numbers are calculated, there is a comparison to see whether census counts or adjusted counts are closer to the target numbers, which are assumed to be "truth." There is still an issue of what is the appropriate comparison between census, adjusted and target numbers. Should it be a simple difference? If so, how are pluses and minuses handled? Should it be the square of the differences, which avoids the problem of pluses and minuses but overemphasizes states (or other areas of interest) with big differences. Or should it be some kind of weighted squared difference to avoid the over-effect of big states but to still reflect some of the differences in state size?

The Committee could come to no consensus on these difficult questions. Therefore, the Committee ran a variety of loss functions. These were a combination of:

-Various methods of allocating the bias to target numbers

-With and without correlation bias

-Absolute and squared error as well as variations of those to take account of variation in state (or other area of interest) size.

Even with these various loss functions, there was still another important question. Do you only look at the aggregate loss over all areas of interest (example, all states), or do you look at individual losses? This question was discussed with the Panel of Experts. The Panel felt that a simple count of "winners" and "losers" was inappropriate. One suggestion was to use a Pitman nearness measure. Time prevented that kind of analysis. In the absence of this measure, the Committee continued its original intent to examine aggregate loss. The Panel supported analysis of aggregate loss. In doing aggregate loss analysis, the Committee heeded the advice of the Panel of Experts who strongly recommended that loss function analysis be viewed only as a tool and not an exact decision mechanism

In examining total loss over a set of areas (like all states), there was a question about whether the difference in aggregate loss between the census and adjusted base counts was a real difference or only due to random error. The Census Bureau had developed a statistical hypothesis test to try to answer that question. The Panel of Experts reviewed this work as well. In particular, the representative from Statistics Canada, who face the same problem, commented on the proposed hypothesis test. That expert warned that in effect we were not doing a standard hypothesis test, but rather we would be making a decision on which set of estimates to use based on the results of the test. If we continued with the standard test, we could be making mistakes about what level of significance to use. The most appropriate level might very well be larger than the 10% level of significance the Committee chose to use. Because of the lateness of the suggestion, time prevented us from completely examining the alternative hypothesis test approach. Hence, the Committee used, with caution, the significance level of standard hypothesis test results.

In summary, using hypothesis tests with 10% significance, loss function analysis excluding correlation bias does not support adjustment. Using hypothesis tests with 10% significance and including correlation bias, all but one of the loss function analyses favors adjustment at the state level when examining aggregate loss[11]. The Committee tended to accept these findings keeping in mind the numerous caveats mentioned above.

FOURTH ISSUE: DOES THE HOMOGENEITY ASSUMPTION HOLD?

> Summary: Just as in July 1991, the results on whether the homogeneity assumption holds are inconclusive. The new research used to examine the homogeneity assumption (called artificial population analysis) indicates that the assumption does not hold when the bias in the estimate gets to be about 25% or higher. Since the bias in the Post Enumeration Survey (PES) estimate as currently measured is 22% to 45%, the Committee was concerned.

An integral part of the PES/DSE system is to assume that everyone within a post-stratum has approximately the same probability of being counted in the PES. This is often referred to as having the same "capture probability." As discussed in the part of the third issue having to do with correlation bias, failure of this assumption leads to a bias in the DSE. It is also important because of the way the sample is selected and used to make estimates for states, cities, etc. Very few political units, including states, have direct estimates from the PES. That is, the state (or city) was not defined as a universe, and then a sample drawn from it to represent it. Rather, the sample was drawn by region, type of area (large urban area, other urban, rural), race, etc. Therefore, a sample case in Tennessee (for example) also is used in the estimate of undercount for Florida, Georgia, etc. This approach assumes homogeneity. Recognizing the importance of this assumption, the Census Bureau designed a study (labeled P-12) to analyze whether the homogeneity assumption held. The results of P-12 were mixed or inconclusive.

Recognizing this, the Committee asked for more extensive research into the issue of homogeneity. The new research was called artificial population analysis. Basically, items felt to be correlated with undercount were selected. They were called surrogate variables. These items were then scaled to the level of the undercount. For example, the mail return rate of census questionnaires was one of these items. The mail return rate was about 65% while undercount was about 2%. The 65% was scaled to 2%. Then an area that had a mail return rate 5% greater than the national average, got a scaled mail return rate 5% above the national average.

We know mail return rates for every area in the country. Using the same process used to estimate DSE's we estimated this scaled mail return rate. In effect, the comparison of the estimated scaled mail return rate to the known scaled mail return rate substitutes for the comparison of estimated undercount with known undercount.

---

[11]Summaries of loss function analysis results can be found in the following C.A.P.E. minutes: C.A.P.E. minutes 5-4-92 Attachment 4; C.A.P.E. minutes 6-1-92 Attachments 9-11; C.A.P.E. minutes 6-9-92 Attachment 5; C.A.P.E. minutes 7-6-92 Attachments 2,3.

Various types of loss function analyses were used to compare the estimated scaled surrogate variables with the actual scaled surrogate variables. If the loss from the estimate was small you could assume that the post-stratification was good and the homogeneity assumption was holding. If the loss was large, there would be cause for concern. In addition, we could examine the number of places (states, cities, etc.) "improved" by adjustment. We could do this kind of analysis for surrogate variables since we know truth (the actual value of the surrogate variable).

Based on artificial population analysis, a first analysis showed similar homogeneity for the 1,392 design as well as the 357 design as well as for a design with only 2 strata. Further analysis showed two problems. One, the surrogate variables did not vary much by post-stratum. Since the assumption was that undercount did vary by post-stratum, there was concern about whether this set of surrogate variables was a good set. Another concern was that the analysis assumed no bias in the surrogate variable estimates and the PES estimates of undercount are biased. Therefore, there was an attempt to find additional surrogate variables as well as to introduce bias into the artificial population analysis. Artificial population analysis was rerun with various levels of constant bias added. The bias in the PES is not constant, but there was no adequate way to introduce variable bias into the artificial population analysis.

The original five surrogate variables were:

-Allocation Rate (The rate at which questions without answers on the census questionnaire had to be allocated a response)
-Percent of population covered by the mail census procedure
-Percent enumerated by mail (mail return rate)
-Substitution rate (The rate at which an entire person's census characteristics had to be created by a computer algorithm)
-Percent of housing units that were multi-unit

The three additional items were:

-Percent in poverty
-Percent unemployed
-A mobility statistic

For states and most large geographic areas, without any bias, artificial population analysis supported the homogeneity assumption assuming that the surrogate variables act like undercount. Once bias is introduced, however, the artificial population analysis shows less and less homogeneity. When bias is 25% of the estimate, the artificial population analysis indicates that there is serious concern that the homogeneity assumption does not hold. Currently, with correlation bias included, the bias in the PES estimate of undercount is 22%. Without correlation bias, the bias is 45% of the estimate. In summary, the Committee could only support the homogeneity assumption with some concern since the level of bias in the PES was close to the point where artificial population analysis shows the homogeneity assumption fails to hold.

FIFTH ISSUE: CAN THE INCONSISTENCY OF PES AND OTHER ESTIMATES BE EXPLAINED?

Summary: Even though there were some points of concern, the Committee is much more comfortable with the consistency of the revised Post Enumeration Survey (PES) estimates and Demographic Analysis (DA) than they were with the July 1991 PES estimates and DA. At the state level, the Committee generally felt the revised PES estimates met their face validity expectations with some individual state exceptions.

As part of the July 1991 decision whether to adjust the 1990 census, there were many concerns about the PES estimates compared to other estimates, mainly Demographic Analysis (DA). In particular, there was concern that the PES estimated a higher population than DA and the fact that the PES estimated about a million more woman than DA. In addition, PES estimates were compared to "best professional judgement" estimates, mainly to see if undercount was being measured by the PES in areas where undercount was expected. This check was called face validity. Face validity checks, though not rigorous, indicated some areas of concern in the PES estimates. For these reasons, the Committee requested additional research to try to investigate the apparent differences.

With regard to DA, the revised PES estimates are now much more consistent. Attachment 7 contains a table summarizing the comparisons. The PES estimate of total population was now lower than the DA estimate, a more expected outcome. The estimated undercount from the PES at the national level was 1.6% compared to an estimate of 1.8% from DA. The PES estimate of women remained higher than DA (an unexpected result), but the difference has been reduced from one million to about 400,000 and was within sampling error. As expected, the PES estimates for Blacks (and in particular, young Black males) were much lower than the DA estimates. This is a result of correlation bias. Even though expected, the Committee was concerned about this problem because there was no method to adequately add these people back into PES estimates.

With regard to face validity checks, there also was now more consistency. Almost all of the changes between the revised PES and the July 1991 PES estimates were in the direction expected by the Committee.

Since intercensal estimates of states are of such importance, the Committee asked for an analysis of revised PES state estimates compared with other information on states to see if there was consistency. Basically, there was consistency with a few exceptions. The exceptions were substantiated by an independent analysis done by one of the Panel of Experts. The Committee was concerned about these exceptions, therefore, they could only conclude that, on average, there would be an improvement using adjusted base counts for states.

In summary, even though there were some points of concern, the Committee was much more comfortable with the consistency of the revised PES estimates and DA than they were with the July 1991 PES estimates and DA. At the state level, the Committee generally felt the revised PES estimates met their face validity expectations with some exceptions.

THE DECISION PROCESS

The decision process that led to the assessment of the Committee contained many parts. By far, the largest part was the year of extensive research and discussion between the Committee and the statistical staff at the Census Bureau. That part of the decision process is summarized in this report and recorded in far more detail in the minutes of the Committee. The decision process culminated with three key discussions. These were a day long meeting with the Panel of Experts, a decision discussion meeting with the Committee, and a decision discussion meeting with the Executive Staff of the Census Bureau. This section of the report summarizes those three meetings.

MEETING WITH PANEL OF EXPERTS:

The Census Bureau wanted to have outside review of the additional research it had done since July 1991. The Census Bureau wanted to include some Panel members who had not been too involved in the July 1991 decision in order to get a fresh look. In addition, the Census Bureau considered the outside expert advice it obtained in conjunction with the July 1991 decision. The Panel of Experts was sent materials in advance. In addition, each member was asked to chose two of five key areas on which to concentrate his or her attention. They were, of course, free to comment on any other issue, and as expected, they did. The meeting with the Panel was held on July 14, 1992. In order to place this summary of the Panel meeting in proper context, it is important to understand that the agenda for the Panel was restricted to major problems and that the Census Bureau specifically requested critical review.

In summary, the Panel made comments on the following key points:

1. The Panel thought the additional research done by the Census Bureau was extremely thorough and useful. The Panel took the time to commend the Census Bureau for this effort. They felt this research took the Census Bureau a long way towards being able to adjust at some time, even if not fully at the present.

2. The Panel thought the Census Bureau should only adjust for the geographic areas for which it was comfortable supporting the decision on technical grounds. Even then, there were bound to be some areas that were adversely affected by an adjustment or no adjustment, even though most were improved. The Panel urged the Census Bureau to examine the exceptions and see if they were "seriously" hurt. If so, the Panel recommended the Census Bureau reconsider an adjustment, even if it was technically defensible on average. For areas below the level for which there is technical backing to support adjustment, the decision about whether to adjust was more of a policy issue. The Panel did point out that errors in estimates of population change from the census year to the year of interest could be large, and perhaps larger than errors from adjustment, particularly for small areas.

3. The Panel cautioned that many of the statistical analyses used by the Census Bureau (Loss Function, Total Error Model, etc.) were just tools and not exact decision mechanisms.

4. The Panel would have felt more comfortable if the bias could be removed from the PES estimates before their use in any potential adjustment. The Census Bureau agreed with the concern of the Panel but knew of no adequate methodology to remove the bias by state, city, etc.

In addition, the Panel expressed some concerns:

1. The Panel was quite concerned about the negative values in the fourth cell. The Panel suggested ways to alter the DSE process in order to avoid the negative values.

2. While the Panel recognized the need to do something about correlation bias, they also recognized the potential problems caused by the inability to estimate the components of the bias separately. The Panel was also concerned about the problems with the proposed allocation scheme.

3. The Panel cautioned against loss function analysis where winners and losers were tallied up. Instead, if the intent is to examine individual losses/gains, the Panel recommended a Pitman nearness measure be used.

4. The Panel cautioned against too much reliance on the significance level in the hypothesis test the Census Bureau was planning to use and urged the Census Bureau to consider the implications of the approach to hypothesis testing being studied by Statistics Canada.

5. The Panel cautioned that artificial population analysis, like the P-12 study, was inconclusive about whether the homogeneity assumption held.

6. Some Panel members expressed concern about the extensive use of synthetic estimation in the adjustment process. (Examples: allocating undercount estimates to areas below which there were direct estimates, allocating bias, etc.)

Attachment 8 contains more detail from the meeting with the Panel of Experts.

C.A.P.E. DECISION DISCUSSION

In July 22, 1992, the Committee met with the Director to discuss each member's opinion about the accuracy of adjusted base counts for use in intercensal estimates. Prior to the main part of the meeting, one of the Committee members made a suggestion based on some analysis he had performed. He recommended the Committee consider a composite (50-50) estimate which would be the simple average of the census count and the adjusted base. The reasoning for the suggestion was that we have two estimates of population, both with error. Despite massive research, it is still inconclusive about which is better overall, for all levels of geography. Therefore, an average of the two might make sense. There is precedent for this kind of averaging in other Census Bureau work. Despite the lateness of the suggestion, the Committee members were asked to comment on the new proposal.

To help in the overall discussion about whether to adjust the base for intercensal estimates, there was a list of key uses and issues of intercensal estimates. Committee members were asked to tie their opinions about potential improved accuracy to the uses of the estimates and geographic level. The list is shown in Attchment 9.

Each Committee member expressed his or her opinion about whether or not the base for intercensal estimates should be adjusted. Though not unanimous, most of the Committee members felt that adjustment of the base should be done at the national and state level. For national and state uses of intercensal estimates, most Committee members felt adjusting the base would make the eventual estimates better on average. There was considerable concern about the states for which it was uncertain whether adjustment would make an improvement. Below the state level, the Committee could not make a recommendation about improvement from adjustment and supported the census counts. In terms of the issue of differential undercount and perception of fairness, the Committee strongly felt that adjustment at the state and national level would satisfy that element. The Committee could come to no agreement on whether an adjustment to the base would improve overall accuracy (accuracy at all levels of geography).

In addition to those summary findings, some other points were raised. These included:

1. No matter what the decision, the Census Bureau needed to examine the existing intercensal estimate challenge system[12]. Regardless of the Census Bureau decision on adjusting the base, a political jurisdiction who feels it was harmed by the Census Bureau decision can and will challenge.

2. Could we adopt the system used in Australia and perhaps Canada? The census is not adjusted, but intercensal estimates are.

---

[12]Currently, there is a challenge system in place that allows jurisdictions to question their intercensal estimates. The evidence supplied by the jurisdiction is reviewed by Census Bureau staff. The staff selected are not involved in the intercensal estimate operations. If the challenge is accepted, the intercensal estimate is changed.

3. No matter what the decision on adjustment of the base for intercensal estimates, the reliance on the current DSE system should be examined. Some of the problems with it might never be solved. (See the final section of this report-FUTURE)

The meeting closed with a discussion of the 50-50 composite suggestion. Only a minority of the Committee favored the 50-50 composite as a first choice, although many of the Committee members thought the composite could be a possible acceptable alternative. During the discussion, several pros and cons of the suggestion were listed.

PROS:

1. It would produce estimates that are additive. A procedure following the Committee's general consensus of states and higher would not be additive.

2. It is a move in the right direction. (This can also be viewed as a con since it is only a partial correction, even at the national level.)

3. It dampens the effect of noise (bias, error, etc.) in the PES and census.

4. At the substate level, the composite is probably better than the full adjustment.

5. Even with an adjustment, there would still be a benefit for respondents to take the effort to be counted in the future, because any potential adjustment based on the 50-50 composite method would only be a partial correction.

6. Analysis done by one Committee member showed that hypothesis test results at the state level were much more favorable to the composite estimate than to the full adjustment, even without including correlation bias.

CONS:

1. It is not as good an estimate at the national level as at the adjusted base, but it is probably a better estimate than an estimate with a fully adjusted base for substate levels. Substate improvement is at the expense of state and national estimates.

2. The two estimates (the DSE and the census) are not independent.

3. It was too late to fully examine the technical merits of the composite.

4. It is only half a solution to differential undercount.

5. It looks like a compromise or even like a "cop-out."

6. Why 50-50? 60-40 or some other combination might be better, and there is no way to know.

EXECUTIVE STAFF DECISION DISCUSSION

Following the Committee discussion, the Executive Staff of the Census Bureau met to give their views. Basically, the Executive Staff concentrated on policy concerns since the Committee had discussed the technical issues. The Executive Staff did not make a recommendation on whether or not to adjust the base for intercensal estimates, but rather raised some issues. The following points were raised at the Executive Staff meeting:

1. It is very important to make sure that people understand that the decision on whether to adjust the base for intercensal estimates is different from the decision whether to adjust the full census. Even if there is a decision to adjust the base for intercensal estimates, there is no intention to adjust the 1990 census because research shows insufficient technical justification.

2. The Census Bureau should do what it thinks it can support based on statistical science.

3. The Census Bureau should consider the advice of users, but should not be forced into a decision because of pressure from users.

4. The Census Bureau should consider the effect of the decision on the public and in particular on its respondents.

5. The 50-50 composite suggestion looks arbitrary.

6. The adjustment issue is so complex, there is probably no single intellectually coherent solution. Most likely, none of the available options is fully consistent with the current research. Also, no matter what the decision, some people will not be satisfied.

On balance, the Executive Staff felt very strongly that there should be technical support for the eventual decision. The Executive Staff recognized that many issues, some of them nontechnical, would need to be balanced in making the final choice. Even so, it is very important for the Census Bureau to be confident about the technical support for the decision it chooses. Not only would the Census Bureau have to defend any decision, but the professionalism of the agency can be questioned if the Census Bureau cannot stand behind its decision on statistical grounds.

FUTURE

Regardless of the choice about whether to adjust the base for intercensal estimates, there were several concerns about the future raised during the final discussions. Generally, it was felt that the problem of differential coverage will continue in the future. Therefore, there were strong recommendations that research in the area of differential undercount should continue as input into the design of the year 2000 census. In particular, the following points were made.

1. The Census Bureau should examine alternatives to the Dual System Estimation process used in 1990. Some of the problems of that approach may continue despite best efforts, meaning that a full adjustment based on such a system might never be possible.

2. Even though it might not be statistically efficient, coverage measurement surveys in the future should have samples and estimation systems that produce direct estimates for key political areas (like states).

3. The Committee process was very successful and could be a good model for the future. Examples of the benefits included sufficient time, timely senior staff input, clear goals, etc.

4. Any proposed undercount estimation/adjustment scheme must be simple. It must be simple enough so the technical aspects can be evaluated and it must be simple enough so it can be explained, even to those without extensive statistical knowledge.

5. Methods of incorporating coverage measurement into the census process should be examined.

6. A system that produces one set of counts rather than unadjusted and adjusted counts is definitely preferred.

Attachment 1: List of

COMMITTEE ON ADJUSTMENT OF POSTCENSAL ESTIMATES (CAPE)

MEMBERS

Dr. Barbara Everitt Bryant	Director
Mr. C. L. Kincannon	Deputy Director
Mr. William Butz	Associate Director
Mr. Charles Jones	Associate Director
Dr. Robert Tortora	Associate Director
Mr. Peter Bounpane	Assistant Director
Ms. Paula Schneider	Chief, Population Division
Mr. John Thompson	Chief, Decennial Statistical Studies Division
Dr. Robert Fay	Senior Mathematical Statistician
Dr. Howard Hogan	Statistical Research Division
Dr. John Long	Population Division
Dr. Mary Mulry	Decennial Statistical Studies Division
Dr. Gregory Robinson	Population Division
Mr. Michael Batutis	Population Division
Mr. David Whitford	Decennial Management Division

Attachment 2

LIST OF MEMBERS OF PANEL OF EXPERTS WHO ATTENDED THE MEETING WITH THE CENSUS BUREAU

Mr. Don Royce
Senior Methodologist
Statistics Canada Social Survey Methods Division

Mr. Wesley Schaible
Associate Commissioner
Office of Research and Evaluation
Bureau of Labor Statistics

Dr. Fritz Scheuren
Director, Statistics of Income Division
Internal Revenue Service

Dr. Bruce Spencer
Department Head
Statistics Department
Northwestern University

Dr. Theresa A. Sullivan
Chair and Professor for the Department of Sociology
University of Texas at Austin

Dr. James Trussell
Associate Dean of Woodrow Wilson School and Director of the Office of Research
Princeton University

Mr. Joseph Waksberg
Chairman of the Board
WESTAT

Dr. Tommy Wright
Research Staff Member
Oak Ridge National Laboratory

Dr. Donald Ylvisaker
Director for the Division of Statistics, Mathematics Department
University of California

Dr. Alan Zaslavsky
Assistant Professor
Statistics Department
Harvard University

ATTACHMENT 3A: PES ESTIMATES OF UNDERCOUNT BY RACE AND SEX
JULY, 1992

Table 1 Table of PES Estimates for Selected Race/Origin/Sex Groups

		JULY, 1991 Original PES		JANUARY, 1992 Revised PES		JULY, 1992 357 PES	
Race/Hispanic/Sex	Census	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Total	248709873	253979141	472946.472	252959473	461310.829	252712821	489754.595
Male	121239418	124249093	245445.426	123648997	238663.637	123623143	273518.304
Female	127470455	129730048	246737.086	129310476	241383.831	129089678	254912.175
Black	29986060	31505838	95559.460	31295058	93635.743	31377094	167925.028
Male	14170151	14974382	49052.934	14857391	47952.832	14900868	82912.806
Female	15815909	16531456	52914.183	16437667	51898.230	16476225	96609.126
Non-Black	218723813	222473303	424675.175	221664415	414933.642	221335728	453076.281
Male	107069267	109274711	222153.799	108791606	216160.510	108722274	249791.220
Female	111654546	113198592	220800.163	112872809	216539.374	112613453	239423.186
Asian or Pacific Islander	7273662	7504906	36264.289	7485602	36157.768	7447371	102828.516
Male	3558038	3688436	19879.800	3674532	19946.424	3684895	60817.829
Female	3715624	3816470	18469.115	3811069	18435.209	3762476	57240.421
American Indian	1878285	1976890	21726.014	1970537	21588.870	2051976	26259.820
Male	926056	980874	1[illegible]12.232	977738	1130[illegible].066	1020059	13248.050
Female	952229	996016	10612.782	992799	10467.531	1031917	13252.478
Hispanic	22354059	23590274	103458.969	23471101	102033.476	23521183	180090.423
Male	11388059	12086513	57498.441	12008888	56356.003	12052241	114778.144
Female	10966000	11503761	52275.143	11462214	52082.441	11468942	84750.443

Table 2 Table of Undercount Rates for Selected Race/Origin/Sex Groups

		Original PES		Revised PES		357 PES	
Race/Hispanic/Sex	Census	UC Rt	SE(UC Rt)	UC Rt	SE(UC Rt)	UC Rt	SE(UC Rt)
Total	248709873	2.075	0.182	1.680	0.179	1.584	0.191
Male	121239418	2.422	0.193	1.949	0.189	1.928	0.217
Female	127470455	1.742	0.187	1.423	0.184	1.254	0.195
Black	29986060	4.824	0.289	4.183	0.287	4.433	0.511
Male	14170151	5.371	0.310	4.626	0.308	4.904	0.529
Female	15815909	4.328	0.306	3.783	0.304	4.008	0.563
Non-Black	218723813	1.685	0.188	1.327	0.185	1.180	0.202
Male	107069267	2.018	0.199	1.583	0.196	1.520	0.226
Female	111654546	1.364	0.192	1.079	0.190	0.852	0.211
Asian or Pacific Islander	7273662	3.081	0.468	2.831	0.469	2.332	1.349
Male	3558038	3.535	0.520	3.170	0.526	3.443	1.594
Female	3715624	2.642	0.471	2.504	0.472	1.245	1.502
American Indian	1878285	4.988	1.044	4.682	1.044	4.520	1.222
Male	926056	5.589	1.089	5.286	1.095	5.183	1.231
Female	952229	4.396	1.019	4.086	1.011	3.864	1.235
Hispanic	22354059	5.240	0.416	4.759	0.414	4.962	0.728
Male	11388059	5.779	0.448	5.170	0.445	5.511	0.900
Female	10966000	4.675	0.433	4.329	0.435	4.385	0.707

Note: Due to the nature of the data used to compute these counts for the 357 poststrata PES design, the American Indian counts in both Table 1 and Table 2 above include Eskimos and Aleuts for the 357 PES. The census count used for this group was 1,959,234. T census counts used to compute the original PES counts and the revised PES counts are shown in the tables.

ATTACHMENT 3B: REVISED PES ESTIMATES OF UNDERCOUNT BY AGE-RACE-SEX
JULY, 1992

Table 1 PES Estimates for Selected Race/Origin/Sex Groups for the 0 to 17 Age Group
(357 Poststrata PES Design)

JULY, 1992

Race/Origin/Sex Group	Census	357 PES Estimate	Std. Error	Undercount Rate	Standard Error
Total	63604432	65695382	191195.568	3.183	0.282
Male	32584278	33649795	97745.288	3.166	0.281
Female	31020154	32045587	93459.542	3.200	0.282
Black	9584415	10311019	95917.245	7.047	0.865
Male	4849497	5215800	48390.736	7.023	0.863
Female	4734918	5095218	47527.287	7.071	0.867
Non-Black	54020017	55384363	172047.616	2.463	0.303
Male	27734781	28433994	88325.776	2.459	0.303
Female	26285236	26950369	83724.989	2.468	0.303
Asian or Pacific Islander	2083387	2152880	46537.029	3.228	2.092
Male	1063264	1099038	23792.412	3.255	2.094
Female	1020123	1053842	22745.817	3.200	2.089
American Indian, Eskimo, or Aleut	696967	742996	12481.466	6.195	1.576
Male	354875	378205	6315.004	6.169	1.567
Female	342092	364791	6166.491	6.222	1.585
Hispanic	7757500	8164834	77292.661	4.989	0.899
Male	3971164	4179630	39551.088	4.988	0.899
Female	3786336	3985204	37742.086	4.990	0.900

Table 2 PES Estimates for Selected Race/Origin/Sex Groups for the 18 to 29 Age Group
(357 Poststrata PES Design)

JULY, 1992

Race/Origin/Sex Group	Census	357 PES Estimate	Std. Error	Undercount Rate	Standard Error
Total	48050811	49530134	192936.681	2.987	0.378
Male	24312055	25105216	129869.843	3.159	0.501
Female	23738756	24424918	113605.768	2.809	0.452
Black	6419397	6727151	60784.870	4.575	0.862
Male	3110320	3225832	38478.198	3.581	1.150
Female	3309077	3501319	41388.086	5.491	1.117
Non-Black	41631414	42802983	174778.637	2.737	0.397
Male	21201735	21879384	121313.350	3.097	0.537
Female	20429679	20923599	102738.356	2.361	0.479
Asian or Pacific Islander	1581231	1686549	47226.618	6.245	2.625
Male	802067	893983	35821.446	10.282	3.595
Female	779164	792566	31415.861	1.691	3.897
American Indian, Eskimo, or Aleut	414071	441408	7298.043	6.193	1.551
Male	210263	224725	4083.000	6.435	1.700
Female	203808	216683	3782.708	5.942	1.642
Hispanic	5525130	5903999	83906.191	6.417	1.330
Male	2984897	3207779	67903.944	6.948	1.970
Female	2540233	2696220	31412.026	5.785	1.098

ATTACHMENT 3B: REVISED PES ESTIMATES OF UNDERCOUNT BY AGE-RACE-SEX
JULY, 1992

Table 3 PES Estimates for Selected Race/Origin/Sex Groups for the 30 to 49 Age Group
(357 Poststrata PES Design)

Race/Origin/Sex Group	Census	July, 1992 357 PES Estimate	Std. Error	Undercount Rate	Standard Error
Total	73314363	74327349	178380.748	1.363	0.237
Male	36281757	36965692	114336.225	1.850	0.304
Female	37032606	37361657	94874.030	0.881	0.252
Black	8300318	8705762	57437.333	4.657	0.629
Male	3841762	4099633	38014.164	6.290	0.869
Female	4458556	4606129	31219.727	3.204	0.656
Non-Black	65014045	65621588	168451.681	0.926	0.254
Male	32439995	32866059	106016.209	1.296	0.318
Female	32574050	32755528	90532.426	0.554	0.275
Asian or Pacific Islander	2373785	2396349	35297.064	0.942	1.459
Male	1128527	1127567	23875.089	-0.085	2.119
Female	1245258	1268782	19001.048	1.854	1.470
American Indian, Eskimo, or Aleut	543821	560400	5746.845	2.958	0.995
Male	263525	276134	2812.700	4.566	0.972
Female	280296	284266	3232.422	1.397	1.121
Hispanic	5961207	6271153	61500.742	4.942	0.932
Male	3029043	3225477	40130.964	6.090	1.168
Female	2932164	3045676	33430.513	3.727	1.057

Table 4 PES Estimates for Selected Race/Origin/Sex Groups for the 50 and Older Age Group
(357 Poststrata PES Design)

Race/Origin/Sex Group	Census	July, 1992 357 PES Estimate	Std. Error	Undercount Rate	Standard Error
Total	63740267	63159956	164191.819	-0.919	0.262
Male	28061328	27902440	91400.020	-0.569	0.329
Female	35678939	35257516	98575.330	-1.195	0.283
Black	5681930	5633162	34874.194	-0.866	0.624
Male	2368572	2359603	22227.003	-0.380	0.946
Female	3313358	3273559	19516.989	-1.216	0.603
Non-Black	58058337	57526794	159823.396	-0.924	0.280
Male	25692756	25542837	89232.535	-0.587	0.351
Female	32365581	31983957	96067.229	-1.193	0.304
Asian or Pacific Islander	1235259	1211593	20586.691	-1.953	1.732
Male	564180	564307	7192.919	0.023	1.274
Female	671079	647286	18017.833	-3.676	2.886
American Indian, Eskimo, or Aleut	304375	307172	3091.413	0.911	0.997
Male	138523	140996	1832.019	1.754	1.277
Female	165852	166176	1554.022	0.195	0.933
Hispanic	3110222	3181198	45726.253	2.231	1.405
Male	1402955	1439356	27996.289	2.529	1.896
Female	1707267	1741842	32679.612	1.985	1.839

ATTACHMENT 4:

JULY, 1992

State Level Estimates and Estimated Undercount Rates

State	1990 Census	Original PES July 1991 Estimate	UC Rt	SE(UCRt)	357 PES July 1992 Estimate	UC Rt	SE(UCRt)
01 Alabama	4040587	4146133	2.546	0.383	4113119	1.763	0.316
02 Alaska	550043	560727	1.905	0.437	561255	1.998	0.364
04 Arizona	3665228	3790186	3.297	0.466	3754297	2.373	0.455
05 Arkansas	2350725	2402925	2.172	0.417	2392291	1.738	0.337
06 California	29760021	30888075	3.652	0.420	30594537	2.728	0.379
08 Colorado	3294394	3376099	2.420	0.470	3363357	2.050	0.383
09 Connecticut	3287116	3305658	0.561	0.556	3308309	0.641	0.406
10 Delaware	666168	686661	2.984	0.437	678372	1.799	0.377
11 District of Columbia	606900	638747	4.986	0.517	628309	3.407	0.901
12 Florida	12937926	13277708	2.559	0.384	13196855	1.962	0.390
13 Georgia	6478216	6632561	2.327	0.368	6618829	2.124	0.331
15 Hawaii	1108229	1136417	2.480	0.537	1129162	1.854	0.808
16 Idaho	1006749	1035271	2.755	0.501	1029213	2.183	0.434
17 Illinois	11430602	11592305	1.395	0.352	11544433	0.986	0.358
18 Indiana	5544159	5585918	0.748	0.370	5572239	0.504	0.399
19 Iowa	2776755	2807238	1.086	0.455	2788378	0.417	0.404
20 Kansas	2477574	2506427	1.151	0.353	2494762	0.689	0.350
21 Kentucky	3685296	3767824	2.190	0.418	3745662	1.612	0.370
22 Louisiana	4219973	4332297	2.593	0.366	4313516	2.169	0.339
23 Maine	1227928	1240076	0.980	0.611	1237124	0.743	0.562
24 Maryland	4781468	4868990	1.798	0.444	4882324	2.066	0.418
25 Massachusetts	6016425	6039315	0.379	0.548	6045161	0.475	0.485
26 Michigan	9295297	9403964	1.156	0.368	9361331	0.705	0.371
27 Minnesota	4375099	4419180	0.998	0.355	4394680	0.446	0.380
28 Mississippi	2573216	2632412	2.249	0.397	2628899	2.118	0.434
29 Missouri	5117073	5184411	1.299	0.352	5149052	0.621	0.363
30 Montana	799065	822092	2.801	0.514	818305	2.351	0.492
31 Nebraska	1578385	1594894	1.035	0.380	1588698	0.649	0.366
32 Nevada	1201833	1231620	2.419	0.469	1230675	2.344	0.383
33 New Hampshire	1109252	1115972	0.602	0.530	1118610	0.837	0.546
34 New Jersey	7730188	7836174	1.353	0.498	7774411	0.569	0.612
35 New Mexico	1515069	1586489	4.502	0.514	1563123	3.074	0.505
36 New York	17990455	18304414	1.715	0.451	18261955	1.487	0.581
37 North Carolina	6628637	6814693	2.730	0.363	6753175	1.844	0.347
38 North Dakota	638800	647837	1.395	0.463	643042	0.660	0.502
39 Ohio	10847115	10933439	0.790	0.354	10921925	0.685	0.360
40 Oklahoma	3145585	3213646	2.118	0.386	3202730	1.784	0.338
41 Oregon	2842321	2898058	1.923	0.445	2896147	1.859	0.401
42 Pennsylvania	11881643	11956891	0.629	0.477	11916630	0.294	0.483
44 Rhode Island	1003464	1006150	0.267	0.556	1004811	0.134	0.590
45 South Carolina	3486703	3589808	2.872	0.407	3558918	2.029	0.362
46 South Dakota	696004	706954	1.549	0.494	702878	0.978	0.548
47 Tennessee	4877185	5012173	2.693	0.386	4963686	1.743	0.344
48 Texas	16986510	17550747	3.215	0.378	17469248	2.763	0.395
49 Utah	1722850	1757423	1.967	0.537	1753121	1.727	0.497
50 Vermont	562758	570651	1.383	0.709	569091	1.113	0.765
51 Virginia	6187358	6352705	2.603	0.351	6313620	2.000	0.353
53 Washington	4866692	4986607	2.405	0.433	4957987	1.841	0.437
54 West Virginia	1793477	1842267	2.648	0.436	1819004	1.403	0.430
55 Wisconsin	4891769	4923844	0.651	0.369	4921997	0.614	0.397
56 Wyoming	453588	466067	2.678	0.481	463569	2.153	0.416
United States Totals	248709873	253979140	2.075	0.182	252712822	1.584	0.191

UC Rt	Undercount Rate as estimated from the PES.
SE(UCRt)	The sampling error of the estimated undercount rate.

ATTACHMENT 5: THE 357 POSTSTRATUM DESIGN FOR POSTCENSAL ESTIMATION--JULY, 1992

The following page defines the 51 poststrata groups and seven age sex groups used to poststratify the Post-Enumeration Survey (PES). These were used to develop dual system estimates for use in the postcensal estimation program. Cross classification of the 51 poststrata groups with the seven age sex groups yields 357 poststrata cells for which dual system estimates have been developed.

The following rough definitions are used:

"Urbanized area 250,000+" means that the PES sample block was part of an Urbanized Area the total population size of which was greater than 250,000.

"Other-urban" refers to all PES blocks that were part of an Urbanized Area not greater than 250,000 or were part of an other urban place.

"Non-urban" means all rural areas and other areas not falling into the above categories.

"Owner/Non-Owner" is determined from the tenure variable on the PES questionnaire. All persons in group quarters are non-owners by definition.

"Asian and Pacific Islander" refers to all people who report themselves as being Asian and Pacific Islander. This group is not restricted to the West or Mid Atlantic as it was in the July, 1991 estimates. Asians and Pacific Islanders of Hispanic origin are included here.

"American Indians on Reservations" include American Indians living on reservations and Tribal Trust Lands. All other concepts (Black, Non-black Hispanic, etc.) are defined as in the census.

"North East" states are Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont, New Jersey, New York, and Pennsylvania.

"South" states include Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia, Alabama, Kentucky, Mississippi, Tennessee, Arkansas, Louisiana, Oklahoma, and Texas.

"Midwest" states are Illinois, Indiana, Michigan, Ohio, Wisconsin, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota.

"West" states include Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming, Alaska, California, Hawaii, Oregon, and Washington.

Revised Poststratification for Postcensal Estimation (357 Design) - 1992

	North East	South	Mid West	West
Non-Hispanic White & Other				
Owner				
Urbanized Areas 250,000 +	1	2	3	4
Other Urban	5	6	7	8
Non-Urban	9	10	11	12
Renter				
Urbanized Areas 250,000 +	13	14	15	16
Other Urban	17	18	19	20
Non-Urban	21	22	23	24
Black				
Owner				
Urbanized Areas 250,000 +	25	26	27	28
Other Urban	29*			
Non-Urban	30*			
Renter				
Urbanized Areas 250,000 +	31	32	33	34
Other Urban	35*			
Non-Urban	36*			
Non-Black Hispanic				
Owner				
Urbanized Areas 250,000 +	37	38	39	40
Other Urban	41*			
Non-Urban	42*			
Renter				
Urbanized Areas 250,000 +	43	44	45	46
Other Urban	47*			
Non-Urban	48*			
Asian & Pacific Islander				
Owner	49*			
Renter	50*			
American Indians on Reservations	51*			

* Indicates that the group is combined across all regions.

Age-Sex Groups

	Males	Females
0 to 17	a	
18 to 29	b	e
30 to 49	c	f
50 and Over	d	g

ATTACHMENT 6:

Total Error of the Net Undercount Rate Assuming No Correlation Bias and Synthetic Estimation of Net Component Errors

JULY, 1992

Evaluation Poststratum	$\hat{U}$*	$\hat{B}(\hat{U})$	St. Dev. $\hat{B}(\hat{U})$	Total St. Dev.	95% Interval
Non-Hispanic White and Other, Owner					
Urban 250k+	-0.50	0.32	0.99	1.06	(-2.95, 1.31)
Other Urban	0.11	0.21	0.25	0.34	(-0.79, 0.59)
Non-Urban	-0.22	0.86	0.87	1.00	(-3.07, 0.92)
Non-Hispanic White and Other, Non-Owner					
Urban 250k+	2.33	-0.06	0.60	0.96	(0.47, 4.32)
Other Urban	2.92	1.70	0.82	1.13	(-1.03, 3.47)
Non-Urban	5.30	0.47	0.74	1.35	(2.13, 7.53)
Black, Non-Black Hispanic, Asian and Pacific Islander, Urban 250k+					
Owner	1.33	0.84	0.44	0.67	(-0.86, 1.82)
Non-Owner	7.13	0.80	0.48	0.94	(4.44, 8.21)
Black, Non-Black Hispanic, Asian and Pacific Islander, Other Urban & Non-Urban					
Owner	2.07	2.38	0.90	1.25	(-2.81, 2.18)
Non-Owner	6.44	3.98	0.94	1.63	(-0.80, 5.72)
National	1.61	0.73	0.30	0.36	(0.17, 1.60)

*Based on PES population only.

Total Error of the Net Undercount Rate Assuming Synthetic Estimation of Net Component Errors

Evaluation Poststratum	$\hat{U}$*	$\hat{B}(\hat{U})$	St. Dev. $\hat{B}(\hat{U})$	Total St. Dev.	95% Interval
Non-Hispanic White and Other, Owner					
Urban 250k+	-0.50	0.31	0.99	1.06	(-2.94, 1.32)
Other Urban	0.11	0.18	0.25	0.34	(-0.76, 0.62)
Non-Urban	-0.22	0.81	0.88	1.00	(-3.03, 0.97)
Non-Hispanic White and Other, Non-Owner					
Urban 250k+	2.33	-0.68	0.76	1.07	(0.87, 5.16)
Other Urban	2.92	1.54	0.84	1.14	(-0.90, 3.65)
Non-Urban	5.30	-0.12	0.90	1.45	(2.52, 8.31)
Black, Non-Black Hispanic, Asian and Pacific Islander, Urban 250k+					
Owner	1.33	0.80	0.45	0.68	(-0.83, 1.87)
Non-Owner	7.13	-1.37	1.30	1.54	(5.42, 11.56)
Black, Non-Black Hispanic, Asian and Pacific Islander, Other Urban & Non-Urban					
Owner	2.07	2.23	0.95	1.28	(-2.71, 2.41)
Non-Owner	6.44	3.55	1.05	1.70	(-0.50, 6.28)
National	1.61	0.35	0.33	0.38	(0.50, 2.03)

*Based on PES population only.

ATTACHMENT 7: COMPARISON OF REVISED PES ESTIMATES VERSUS DA--JULY 1992

Comparison of the Census, Post Enumeration Survey (PES) and Demographic Analysis (DA) Estimates of Population and Percent Net Undercount: 1990

(A positive difference means that the demographic estimate is higher than the PES estimate; a negative difference means that the demographic estimate is lower).

Race, Sex, Age	PES Net Undercount				DA Net Undercount		Difference in DA and PES Percent Net Undercount	
	Original Estimates July 1991		Revised Estimates July 1992					
	Amount (1)	Percent (2)	Amount (3)	Percent (4)	Amount (5)	Percent (6)	Original PES 7=6-2	357 PES 8=6-4
TOTAL	5,269,267	2.07	4,002,947	1.58	4,683,913	1.85	-0.23	0.26
Male	3,009,674	2.42	2,383,724	1.93	3,480,216	2.79	0.37	0.86
Female	2,259,593	1.74	1,619,223	1.25	1,203,697	0.94	-0.81	-0.32
BLACK	1,519,776	4.82	1,391,033	4.43	1,836,272	5.68	0.86	1.25
Male	804,233	5.37	730,717	4.90	1,338,380	8.49	3.12	3.59
Female	715,543	4.33	660,316	4.01	497,892	3.01	-1.32	-1.00
NONBLACK	3,749,491	1.69	2,611,914	1.18	2,847,641	1.29	-0.40	0.11
Male	2,205,441	2.02	1,653,007	1.52	2,141,836	1.97	-0.05	0.45
Female	1,544,050	1.36	958,907	0.85	705,805	0.63	-0.73	-0.22

NOTE: Original PES estimates are the July 15, 1991 estimates based on 1392 poststrata and incorporate smoothing; revised PES estimates are the July 1992 estimates based on 357 poststrata, all PES revisions since July 1991, and no smoothing.

ATTACHMENT 8: THE MEETING WITH THE PANEL OF EXPERTS

While the Panel came to no consensus about whether the base for intercensal estimates should be adjusted, the Panel was extremely impressed with the extensive research done by the Census Bureau. The concerns raised by the Panel were not criticisms of the Census Bureau's work, but rather were indications of the difficulty and complexity of the overall issue as well as the fact that some of these problems may never be fully solved. The Panel concentrated its discussion on five areas as requested by the Census Bureau. These were the most difficult problem areas that Census Bureau statisticians had not been able to fully resolve. Not only was the discussion limited to difficult problem areas, but as requested by the Census Bureau, the Panel members were critical and raised concerns. Reading just a list of concerns can lead to an unbalanced view of what Panel members felt about the adjustment issue in general. Therefore, the parameters under which the Panel operated should be kept in mind in order to put the following more detailed discussion of Panel concerns in proper perspective.

FIRST AREA: TOTAL ERROR MODEL INCLUDING CORRELATION BIAS

During this discussion the Panel mentioned that it didn't see an easy alternative to the current method of treating correlation bias, but Panel members were uneasy about certain aspects of it. For one, the Panel was quite concerned about the negative fourth cells. In addition, there was concern that we weren't estimating the level of the bias properly. In particular, one Panel member felt we should consider comparing the unbiased PES estimates (taking out the bias) to DA in order to estimate the level of correlation bias. Another panel member expressed serious concern that the Census Bureau assumed all correlation bias was male. This panel member pointed to his research to show that there also are problems of differing capture probabilities in the female population. Currently, the Census Bureau's treatment of correlation bias assumes that doesn't occur. It was also during this discussion that most of the Panel recommended that the Census Bureau try to remove the bias from the PES estimates before making any adjustment. Another panel member went through the PES/DSE process in some detail with an emphasis on whether or not it was understandable to an average person and whether or not it was creditable. He pointed out several parts of the process that were of concern to him particularly the extensive use of synthetic estimation. He also cautioned that if new research between July 1991 and the present uncovered new findings, then he wouldn't be surprised to see additional research after July 1992 turn up new results and new estimates of undercount. Another Panel member strongly desired that total error be broken out separately by persons of Hispanic ethnicity. This section of the meeting concluded with a discussion of the problem of inconsistent race classification between systems (example: PES and DA), which the Panel felt was a significant issue that needed further research.

SECOND AREA: LOSS FUNCTION ANALYSIS

This part of the meeting was quite technical, with a review of the various loss functions under consideration. Most of the Panel advised against counting up winners and losers (For example: states that gained or lost in a loss function analysis done on states). Instead one Panel member recommended a Pitman nearness measure which he uses when faced with this kind of problem. Then, there was a discussion of aggregate loss. The Panel pointed out that decisions on aggregate loss may make sense statistically, but that the "losing" political areas might have a problem. Also, it was during this discussion that the Panel made a recommendation that the results of loss function analysis be used with caution. Loss function analysis is a tool, depends on personal standards of judgement, and is not an exact decision mechanism. It also was during this discussion that the Panel reiterated a theme they raised in the first topic. Panel members were concerned that there is too much confusion about the undercount/adjustment issue by the "person on the street." The Panel recommended that the Census Bureau try to alleviate that in the future. Finally, there was a discussion about the large number of states for which it doesn't matter much whether or not there is an adjustment. Both sides of the case were discussed. If so, why bother to adjust?; or if so, adjust all states in order to correct a problem in a few states and the error in most other states won't be too bad. This discussion ended with another theme heard often. The total error model is a good tool to try alternative assumptions. It is not an exact decision mechanism.

THIRD AREA: HYPOTHESIS TESTS

The Census Bureau had recognized the limitations of loss function analysis. In particular, once you had two losses to compare, was the difference between them a "real" difference, or could it be attributable solely to chance since these were sample estimates. To help answer that question, the Census Bureau planned some statistical hypothesis tests. The Panel was asked to review the Census Bureau plans.

This part of the discussion was led by the expert from Statistics Canada, since Statistics Canada was faced with a similar problem. The discussion was extremely technical. Before getting to the issue of the hypothesis test, the Panel member cautioned that several key questions had to be answered, and they all had an effect on the eventual hypothesis test. These questions included:

What is the quantity of interest? (Total population, population share, etc.)

Which Loss Function would be used?

How accurate are your target numbers?

How do you account for error in estimating the target numbers?

The bulk of the discussion centered about the technical performance of the hypothesis test assuming the above questions had been answered satisfactorily. Basically, the Panel pointed out that we were not simply dealing with a standard hypothesis test. Instead, we planned to use one of the set of estimates based on the results of the hypothesis test. Under those conditions, a model could be developed to examine the true level of risk for the hypothesis test. At present, Statistics Canada had developed such an approach. The Panel member urged the Census Bureau to take this finding into account in the significance level of the Census Bureau's proposed hypothesis test. During this part of the discussion, this panel member warned that if there is a high positive bias in the estimate of undercount, then the hypothesis test can be misleading, and in fact, adjustment can be very problematic when the estimate of undercount has a large bias. Also, it was pointed out that Statistics Canada feels its estimates of undercount at the province level are adequate for use in adjusting intercensal estimates, but not at sub-province level. Whether or not to adjust below the Province level will be more a policy call than a technical decision. Finally, it was during this part of the meeting that the Panel repeated its recommendation that if estimates of bias are good enough for use in determining target numbers for loss function analysis, then they should be removed from the PES estimates before any potential adjustment.

FOURTH AREA: ARTIFICIAL POPULATION ANALYSIS

Because of the way the PES/DSE system operates, the homogeneity assumption is a key one. In conjunction with the July 1991 decision, the Census Bureau studied homogeneity and recorded the results in study called P-12. Since the homogeneity assumption was so key, the Census Bureau undertook additional work in a study called Artificial Population Analysis. The Panel was asked to examine various aspects of the analysis. The Panel member who did part of the P-12 study led the discussion. The Panel member started with a brief review of study P-12 which he characterized as inconclusive. In reviewing the artificial population analysis, he thought the Census Bureau had taken a major additional step to try to investigate the issue, but he still felt the results were inconclusive. In his opinion, only two of the eight surrogate variables considered by the Census Bureau were associated enough with undercount to be considered. (Percent enumerated by mail and substitution rate.) He wondered if there were better alternative surrogate variables. The Panel also expressed some concern about the constant scaling of the surrogate variables to undercount. Variable scaling might be preferred. Likewise, the Panel was concerned about the constant introduction of bias into the artificial population analysis. Once again, variable bias would be preferred. Even so, the Panel was concerned that artificial population analysis showed failure of the homogeneity assumption when the constant bias was 25% or greater. One panel member did some work on his own. From that study, he concluded that by using substitution rate, adjustment looks better. Using poverty, the results are mixed. And, using unemployment rate, the census looks better. This kind of analysis supports the conclusion that even with all the new research, the results are inconclusive. This panel member felt that a considerable amount of additional work would be needed to get a definitive answer on whether the homogeneity assumption held.

FIFTH AREA: COMPARISON OF PES TO DA

Generally, at the national level, estimates of population from DA are felt to be "better" than estimates from a post-censal survey. Even so, the DA estimates are subject to some error. Before discussing the comparison of the PES and DA, one panel member shared her work on the quality of DA numbers. In addition to the known problems with DA, she pointed out some additional places where the DA estimates could be in error. These included:

1. Over correction for the under-registration of black males. (This error has the effect of overestimating the undercount.)

2. The problem of Mexicans near the border who register the birth in the US, but then return to Mexico to raise the child. (This problem has the effect of overstating the undercount.)

3. Under reporting of infant deaths near the border since the birth certificate can be resold. (This problem overstates the undercount.)

4. Concerns about the consistency and reliability of reporting data on vital statistics forms, especially those done by a third party. (These types of errors might not effect the estimate of total undercount, but would effect the estimates by age-race-sex.)

5. Concern about a change in a person's self perception of race/Hispanic over time. These characteristics could be recorded one way at birth and another at death. (This problem only has an effect on DA estimates of undercount by race/Hispanic.)

Even with these and other problems, there is still general confidence in the DA estimates, particularly at the national level. That is why the Panel was concerned about some inconsistencies between the PES and DA. In particular, one panel member reviewed the Census Bureau work that compared PES estimates by state with DA and other information. She was quite concerned about the states that seemed quite inconsistent. At this point, another panel member indicated that another independent study he had done confirmed the inconsistency in a similar set of states. The Panel discussed the issue and concluded that in an adjustment where there would be overall improvement for states, some states would be adversely affected, even if most were improved and the US average was improved. The Panel strongly recommended that the Census Bureau examine if these exception states were hurt "seriously."

The meeting closed with a brief discussion of the actual mechanism of the intercensal estimate process. During that discussion, there was a question about the accuracy of intercensal estimates. That question couldn't be answered exactly, but there was some summary information provided. Basically, by comparing the estimate in a census year to the census count, you can estimate the error in the estimates over a 10-year period. The following table summarizes the Census Bureau findings.

AREA	LEVEL OF ERROR OVER 10 YEARS[1]
States	1.5 - 2.5%
Places over 50,000	4.0%
Places 5,000 to 50,000	7.0 - 8.0%
Places under 5,000	16.0 - 20.0%

[1]Level of error as measured in previous decades. These error estimates exclude any estimated undercoverage in the census.

ATTACHMENT 9: USES OF INTERCENSAL ESTIMATES AND ISSUES CONSIDERED BY C.A.P.E

Uses of Intercensal Estimates:

1. Survey controls
2. Denominators for per capita Federal statistics
3. Funding programs
    a. State populations either for direct funding or as the first tier in a funding program
    b. Substate areas of 100,000 population or larger
    c. Substate areas below 100,000 population

Other Concerns:

1. National population estimates
2. Differential undercount and the perception of fairness
3. Overall accuracy

ATTACHMENT 10: ESTIMATED UNDERCOUNT/OVERCOUNT FOR 51 POST- STRATA, JULY 1992

Post- Strata Groups	Percent Undercount				Standard Errors			
	North East	South	Mid West	West	NE	S	MW	W
Non-Hispanic White & Other								
Owner								
Urbanized Areas 250,000+	-2.13	0.68	-0.26	0.34	1.08	0.71	0.39	0.65
Other Urban	-1.08	0.52	-0.10	0.62	0.49	0.42	0.40	0.58
Non-Urban	-0.54	0.18	-0.71	0.29	0.70	0.69	1.18	0.69
Non-owner								
Urbanized Areas 250,000+	1.16	2.56	2.33	3.18	1.39	1.48	1.61	1.62
Other Urban	3.41	3.20	1.23	4.49	1.51	1.74	1.09	1.34
Non-Urban	6.52	6.23	2.85	6.08	4.20	1.71	1.51	1.81
Black								
Owner								
Urbanized Areas 250,000+	1.63	2.16	0.81	6.10	1.91	0.90	0.87	1.91
Other Urban		1.34				0.98		
Non-Urban		3.52				1.90		
Non-owner								
Urbanized Areas 250,000+	8.37	6.27	5.99	9.96	1.61	1.90	1.68	2.72
Other Urban		4.15				1.18		
Non-Urban		4.62				5.33		
Non-Black Hispanic								
Owner								
Urbanized Areas 250,000+	0.67	2.53	-4.33	2.89	4.45	0.90	2.58	0.87
Other Urban		0.94				1.64		
Non-Urban		2.73				2.69		
Non-owner								
Urbanized Areas 250,000+	6.72	9.34	6.64	5.91	3.51	2.59	3.26	1.84
Other Urban		6.60						
Non-Urban		15.80				5.01		
Asian and Pacific Islander								
Owner		-1.45						
Non-owner		6.96				2.52		
American Indians on Reservations		12.22				4.73		

Negative numbers in table signify as estimated overcount.

AREA	LEVEL OF ERROR OVER 10 YEARS[13]
States	1.5 - 2.5%
Places over 50,000	4.0%
Places 5,000 to 50,000	7.0 - 8.0%
Places under 5,000	16.0 - 20.0%

ATTACHMENT 9: USES OF INTERCENSAL ESTIMATES AND ISSUES CONSIDERED BY C.A.P.E.

Uses of Intercensal Estimates:

1. Survey controls

2. Denominators for per capita Federal statistics

3. Funding programs

    A. State populations either for direct funding or as the first tier in a funding program

    B. Substate areas of 100,000 population or larger

    C. Substate areas below 100,000 population

Other Concerns:

1. National population estimates

2. Differential undercount and the perception of fairness

3. Overall accuracy

[13] Level of error as measured in previous decades. These error estimates exclude any estimated undercoverage in the census.

Attachment 11: Place Level Estimates and Estimated Undercount Rates (Places with 100,000 or More Population)

State/	Place/	Place Name	1990 Census	Original PES July 1991	Estimated UCRt	SE(UCRt)	357 PES July 1992	Estimated UCRt	SE(UCRt)
01	0185	Birmingham City	265968	278776	4.594	0.504	273918	2.902	0.750
01	0935	Huntsville City	159789	165498	3.450	0.557	162535	1.689	0.587
01	1165	Mobile City	196278	203932	3.753	0.522	201181	2.437	0.619
01	1180	Montgomery City	187106	194786	3.943	0.516	190738	1.904	0.521
02	0140	Anchorage City	226338	231238	2.119	0.671	232174	2.514	0.518
04	0140	Glendale City	148134	151575	2.270	0.663	150735	1.725	0.371
04	0215	Mesa City	288091	296297	2.770	0.583	292643	1.556	0.638
04	0260	Phoenix City	983403	1013566	2.976	0.569	1003800	2.032	0.515
04	0305	Scottsdale City	130069	132778	2.040	0.589	131178	0.846	0.612
04	0360	Tempe City	141865	147232	3.645	0.588	145453	2.467	0.791
04	0380	Tucson City	405390	419413	3.344	0.577	415971	2.544	0.542
05	1195	Little Rock City	175795	181658	3.228	0.496	179875	2.268	0.610
06	0070	Anaheim City	266406	277711	4.071	0.530	273740	2.679	0.538
06	0180	Bakersfield City	174820	179683	2.706	0.574	179398	2.552	0.511
06	0245	Berkeley City	102724	107538	4.477	0.487	106630	3.664	0.712
06	0525	Chula Vista City	135163	140021	3.470	0.584	138715	2.561	0.475
06	0595	Concord City	111348	113121	1.567	0.622	113137	1.582	0.580
06	0880	El Monte City	106209	112288	5.414	0.745	110792	4.137	0.614
06	0935	Esdondido City	108653	112428	3.374	0.533	111040	2.166	0.549
06	1080	Fremont City	173339	177040	2.091	0.584	176094	1.565	0.522
06	1090	Fresno City	354202	369030	4.018	0.497	366527	3.363	0.555
06	1095	Fullerton City	114144	166779	2.256	0.583	116725	2.211	0.514
06	1110	Garden Grove City	143050	146505	2.358	0.572	146412	2.296	0.515
06	1130	Glendale City	180038	183360	1.812	0.584	184515	2.426	0.579
06	1225	Hayward City	111498	115752	3.675	0.566	114720	2.809	0.503
06	1300	Huntington Beach City	181519	183976	1.336	0.632	184639	1.690	0.635
06	1340	Ingelwood City	109602	123350	11.146	0.953	116991	6.316	1.290
06	1347	Irvine City	110330	111773	1.291	0.631	112191	1.659	0.665
06	1610	Long Beach City	429433	450964	4.774	0.466	445925	3.698	0.594
06	1630	Los Angeles City	3485398	3671205	5.061	0.514	3624206	3.830	0.651
06	1790	Modesto City	164730	168273	2.106	0.601	168849	2.440	0.500
06	1849	Moreno Valley City	118779	126583	6.165	0.563	121925	2.580	0.457
06	1970	Oakland City	372242	392769	5.226	0.540	391553	4.932	0.919

06	1990	Oceanside City	128398	132708	3.248	0.586	131771	2.515	0.510
06	2005	Ontario City	133179	141469	5.860	0.577	137458	3.113	0.551
06	2015	Orange City	110658	113020	2.090	0.590	112738	1.845	0.495
06	2050	Oxnard City	142216	148120	3.986	0.581	147164	3.362	0.643
06	2125	Pasadena City	131591	137947	4.608	0.460	136431	3.548	0.582
06	2230	Pomona City	131723	138469	4.872	0.536	137116	3.933	0.693
06	2278	Rancho Cucamonge City	101409	106655	4.919	0.548	103309	1.839	0.485
06	2370	Riverside City	226505	233085	2.823	0.562	232608	2.624	0.492
06	2420	Sacramento City	369365	384466	3.928	0.477	380736	2.987	0.538
06	2435	Salinas City	108777	113243	3.944	0.595	112703	3.484	0.993
06	2450	San Bernardino City	164164	170413	3.667	0.524	170249	3.574	0.577
06	2475	San Diego City	1110549	1156224	3.950	0.476	1143032	2.842	0.527
06	2485	San Fransisco City	723959	756182	4.261	0.504	745573	2.899	0.626
06	2510	San Jose City	782248	814783	3.993	0.520	801296	2.377	0.474
06	2570	Santa Ana City	293742	309907	5.216	0.648	305815	3.948	0.871
06	2583	Santa Clarita City	110642	112528	1.676	0.647	111997	1.210	0.558
06	2615	Santa Rosa City	113313	115042	1.503	0.668	115898	2.231	0.533
06	2702	Simi Valley City	100217	104425	4.030	0.566	102006	1.754	0.449
06	2805	Stockton City	210943	218902	3.636	0.540	218358	3.396	0.600
06	2835	Sunnyvale City	117229	119490	1.892	0.578	119999	2.308	0.610
06	2897	Thousand Oaks City	104352	108398	3.733	0.565	105407	1.001	0.553
06	2910	Torrance City	133107	134632	1.133	0.601	135125	1.494	0.564
06	3000	Vallejo City	109199	113359	3.670	0.550	112178	2.656	0.544
08	0055	Aurora City	222103	227295	2.284	0.673	227110	2:205	0.583
08	0240	Colorado Springs City	281140	289844	3.003	0.572	287033	2.053	0.635
08	0320	Denver City	467610	482714	3.129	0.579	480862	2.756	0.498
08	0760	Lakewood City	126481	128314	1.429	0.680	128094	1.259	0.649
09	001010	Bridgeport Town	141686	143879	1.524	0.857	145631	2.709	1.029
09	001090	Stamford Town	108056	108286	0.212	0.770	109430	1.256	0.461
09	003070	Hartford Town	139739	143285	2.475	0.957	146308	4.490	1.231
09	009075	New Haven Town	130474	132416	1.467	0.844	135057	3.393	0.842
09	009120	Waterbury Town	108961	109092	0.120	0.759	110722	1.591	0.534
11	0005	Washington City	606900	638747	4.986	0.517	628309	3.407	0.901
12	0645	Fort Lauderdale City	149377	153932	2.959	0.490	152687	2.168	0.660
12	0860	Hialeah City	188004	196416	4.283	0.935	197448	4.783	1.621
12	0915	Hollywood City	121697	125104	2.723	0.509	123463	1.431	0.569
12	1003	Jacksonville City (remainder)	635230	658739	3.569	0.462	649437	2.188	0.548

12	1370	Miami City	358548	376424	4.749	0.703	377379	4.990	1.527
12	1600	Orlando City	164693	170303	3.294	0.462	169260	2.698	0.700
12	1900	St. Petersburg City	238629	245561	2.823	0.472	242149	1.454	0.555
12	2070	Tallahassee City	124773	129647	3.759	0.526	127834	2.395	0.816
12	2075	Tampa City	280015	291356	3.893	0.449	287445	2.585	0.627
13	0150	Atlanta City	394017	415204	5.103	0.540	407923	3.409	0.912
13	0660	Columbus City (remainder)	178681	184860	3.343	0.505	182489	2.087	0.554
13	1725	Macon City	106612	110227	3.280	0.542	109027	2.215	0.586
13	2540	Savannah City	137560	142220	3.277	0.531	140538	2.119	0.560
15	0110	Honolulu CDP	365272	382505	4.505	0.803	372146	1.847	0.989
16	0090	Boise City	125738	127612	1.469	0.702	128336	2.024	0.542
17	1051	Chicago City	2783726	2857364	2.577	0.582	2852041	2.395	0.769
17	4590	Peoria City	113504	116740	2.772	0.681	114753	1.089	0.416
17	4965	Rockford City	139426	143232	2.657	0.681	140598	0.834	0.422
17	5480	Springfield City	105227	107883	2.462	0.700	105921	0.655	0.456
18	0775	Evansville City	126272	129192	2.260	0.712	126950	0.534	0.475
18	0825	Fort Wayne City	173072	177949	2.741	0.690	174511	0.824	0.429
18	0905	Gary City	116646	122166	4.518	0.866	119611	2.479	0.719
18	1145	Indianapolis	731327	737483	0.835	0.612	741712	1.400	0.523
18	2375	South Bend City	105511	108564	2.812	0.681	106417	0.851	0.377
19	0670	Cedar Rapids City	108751	110887	1.926	0.648	109199	0.410	0.430
19	1130	DES Moines City	193187	197761	2.313	0.631	194978	0.919	0.506
20	1430	Kansas City	149767	153306	2.309	0.483	151947	1.435	0.494
20	2194	Overland Park City	111790	112871	0.958	0.491	112485	0.618	0.480
20	2795	Topeka City	119883	123028	2.556	0.602	120748	0.716	0.434
20	3040	Wichita City	304011	308747	1.534	0.480	307807	1.233	0.518
21	1160	Lexington Fayette	225366	233157	3.342	0.602	229930	1.985	0.705
21	1230	Louisville City	269063	279912	3.876	0.499	274816	2.094	0.616
22	0095	Baton Rouge City	219531	227504	3.505	0.479	226061	2.889	0.704
22	0956	New Orleans City	496938	514558	3.424	0.486	513936	3.307	0.876
22	1240	Shreveport City	198525	205361	3.329	0.482	203753	2.566	0.633
24	0025	Baltimore City	736014	772082	4.672	0.511	759127	3.045	0.808
25	013090	Springfield City	156983	158023	0.658	0.785	159597	1.638	0.850
25	017130	Lowell City	103439	103118	0.311	0.770	105772	2.206	0.667
25	025005	Boston City	574283	579743	0.942	0.806	590703	2.780	0.784
25	027300	Worcester City	169759	169075	0.405	0.753	171148	0.812	0.816
26	0080	Ann Arbor City	109592	112804	2.847	0.727	111442	1.660	0.522
26	0680	Detroit City	1027974	1064760	3.455	0.622	1056180	2.671	0.727
26	0920	Flint City	140761	146209	3.726	0.703	143923	2.197	0.584

26	1085	Grand Rapids City	189126	194874	2.950	0.666	191834	1.412	0.504
26	1485	Lansing City	127321	131473	3.158	0.684	129424	1.625	0.553
26	1565	Livonia City	100850	101462	0.603	0.527	100700	0.149	0.364
26	2583	Sterling Heights City	117810	118625	0.687	0.514	117955	0.123	0.402
26	2790	Warren City	144864	145814	0.652	0.535	145018	0.106	0.394
27	2585	Minneapolis City	368383	374965	1.755	0.469	374537	1.643	0.605
27	3425	St. Paul City	272235	275845	1.309	0.485	275962	1.351	0.560
28	0615	Jackson City	196637	205662	4.388	0.515	202591	2.939	0.719
29	2125	Independence City	112301	113335	0.912	0.487	112970	0.592	0.493
29	2220	Kansas City	435146	444859	2.183	0.472	441627	1.468	0.516
29	3875	St. Louis City	396685	408263	2.836	0.518	405175	2.096	0.682
29	4075	Springfield City	140494	143438	2.053	0.650	141440	0.669	0.501
31	1425	Lincoln City	191972	196234	2.172	0.660	193365	0.720	0.455
31	1825	Omaha City	335795	340507	1.384	0.476	339436	1.073	0.498
32	0065	Las Vegas City	258295	266308	3.009	0.562	264680	2.412	0.535
32	0090	Reno City	133850	136305	1.801	0.650	137829	2.887	0.670
34	1715	Elizabeth City	110002	111988	1.773	0.740	113626	3.189	1.244
34	2290	Jersey City	228537	236712	3.454	0.681	236914	3.536	0.942
34	2895	Newark City	275221	285923	3.743	0.775	289965	5.085	1.113
34	3115	Paterson City	140891	146967	4.134	0.752	146865	4.068	1.332
35	0015	Albuquerque City	384736	397206	3.139	0.583	393462	2.218	0.480
36	0030	Albany City	101082	103456	2.295	0.692	103108	1.965	0.802
36	0450	Buffalo City	328123	333145	1.508	0.592	334286	1.844	0.726
36	2505	New York City	7322564	7552196	3.041	0.588	7567146	3.232	0.921
36	3100	Rochester City	231636	239832	3.417	0.720	237133	2.318	0.746
36	3565	Syracuse City	163860	167479	2.161	0.683	166653	1.676	0.769
36	4075	Yonkers City	188082	192435	2.262	0.664	190656	1.350	0.852
37	0480	Charlotte City	395934	412466	4.008	0.467	405932	2.463	0.635
37	0750	Durham City	136611	141713	3.600	0.536	139962	2.394	0.712
37	1065	Greensboro City	183521	189851	3.334	0.518	187128	1.928	0.646
37	2020	Raleigh City	207951	215573	3.536	0.520	213485	2.592	0.728
37	2785	Winston-Salem City	143485	148215	3.191	0.513	146388	1.983	0.619
39	0035	Akron City	223019	229527	2.835	0.683	226256	1.431	0.520
39	0865	Cincinnati City	364040	369165	1.388	0.631	372392	2.243	0.719
39	0900	Cleveland City	505616	512581	1.359	0.637	516598	2.126	0.650
39	0960	Columbus City	632910	639303	1.000	0.605	645256	1.913	0.630
39	1110	Dayton City	182044	188260	3.302	0.670	185861	2.054	0.624
39	4265	Toledo City	332943	335164	0.663	0.600	337317	1.297	0.497
40	1815	Oklahoma City	444719	454958	2.251	0.516	454630	2.180	0.548

40	2465	Tulsa City	367302	375358	2.146	0.539	374856	2.015	0.597
41	0360	Eugene City	112669	114413	1.524	0.702	115726	2.641	0.685
41	0905	Portland City	437319	450413	2.907	0.538	445566	1.851	0.659
41	1005	Salem City	107786	109189	1.285	0.652	110240	2.227	0.546
42	0165	Allentown City	105090	105902	0.767	0.627	105216	0.120	0.831
42	3685	Erie City	108718	110075	1.233	0.662	109866	1.045	0.534
42	7180	Philadelphia City	1585577	1606249	1.287	0.609	1608942	1.452	0.742
42	7234	Pittsburgh City	369879	374002	1.102	0.583	373752	1.036	0.728
44	007065	Providence City	160728	161519	0.490	0.777	164304	2.176	0.829
46	1225	Sioux Falls City	100814	102712	1.848	0.658	101208	0.389	0.496
47	0245	Chattanooga City	152466	157807	3.385	0.528	155875	2.187	0.637
47	0760	Knoxville City	165121	170454	3.129	0.587	168582	2.053	0.698
47	0940	Memphis City	610337	640010	4.636	0.498	628329	2.864	0.709
47	1016	Nashville-Davidson (remainder)	488374	508302	4.109	0.519	499383	2.205	0.625
48	0015	Abilene City	106654	109869	2.926	0.515	108885	2.049	0.646
48	0100	Amarillo City	157615	162215	2.836	0.532	160530	1.816	0.577
48	0175	Arlington City	261721	272160	3.836	0.510	269098	2.742	0.608
48	0210	Austin City	465622	480242	3.044	0.501	483156	3.629	0.752
48	0320	Beaumont City	114323	118161	3.248	0.474	116654	1.998	0.500
48	0980	Corpus Christi City	257453	264658	2.722	0.551	267127	3.622	0.798
48	1085	Dallas City	1006877	1057658	4.801	0.508	1043947	3.551	0.727
48	1340	El Paso City	515342	531606	3.059	0.637	538250	4.256	0.964
48	1500	Fort Worth City	447619	467853	4.325	0.490	461686	3.047	0.606
48	1580	Garland City	180650	185940	2.845	0.494	185336	2.528	0.539
48	1975	Houston City	1630553	1715633	4.959	0.542	1697301	3.933	0.777
48	2060	Irving City	155037	162091	4.352	0.530	160622	3.477	0.762
48	2400	Laredo City	122899	127296	3.454	0.793	126611	2.932	1.262
48	2565	Lubbock City	186206	192375	3.207	0.512	190661	2.336	0.688
48	2795	Mesquite City	101484	104448	2.838	0.503	103803	2.234	0.541
48	3200	Pasadena City	119363	123270	3.170	0.588	123539	3.380	0.721
48	3310	Plano City	128713	132377	2.768	0.519	131188	1.887	0.540
48	3745	San Antonio City	935933	964071	2.919	0.561	974099	3.918	0.857
48	4415	Waco City	103590	107015	3.201	0.476	106382	2.624	0.728
49	0870	Salt Lake City	159936	162897	1.818	0.664	163014	1.888	0.721
51	0025	Alexandria City	111183	112748	1.388	0.541	114451	2.856	0.771
51	0242	Chesapeake City	151976	153512	1.001	0.556	155185	2.068	0.509
51	0590	Hampton City	133793	139284	3.942	0.459	137415	2.636	0.617
51	0860	Newport News City	170045	178053	4.498	0.468	175121	2.899	0.689
51	0875	Norfolk City	261229	273457	4.472	0.444	269011	2.893	0.733

51	0990	Portsmouth City	103907	108477	4.213	0.474	106837	2.742	0.695
51	1035	Richmond City	203056	209959	3.288	0.549	208987	2.838	0.817
51	1280	Virginia Beach City	393069	408213	3.710	0.487	402092	2.244	0.558
53	1140	Seattle City	516259	534576	3.427	0.506	528151	2.252	0.670
53	1220	Spokane City	177196	179308	1.178	0.711	179391	1.223	0.739
53	1280	Tacoma City	176664	180714	2.241	0.625	180831	2.304	0.622
55	1475	Madison City	191262	196296	2.565	0.734	193499	1.156	0.504
55	1645	Milwaukee City	628088	635933	1.234	0.601	642860	2.298	0.681

UCRT Undercount Rate as estimatred etc.

Attachment 12: County Level Estimates and Estimated Undercount Rates (counties with 100,000 or more population)

State	County	County Name	1990 Census	Original PES July 1991 Estimate UC Rt SE(UCRt)			357 PES July 1992 Estimate UC Rt SE(UCRt)		
01	015	Calhoun County	116034	119037	2.523	0.466	117856	1.546	0.424
01	073	Jefferson County	651525	673700	3.292	0.423	665329	2.075	0.517
01	089	Madison County	238912	246704	3.158	0.456	242937	1.657	0.478
01	097	Mobile County	378643	390685	3.082	0.417	387137	2.194	0.479
01	101	Montgomery County	209085	217215	3.743	0.481	213105	1.886	0.480
01	103	Morgan County	100043	102781	2.664	0.459	101438	1.375	0.407
01	125	Tuscaloosa County	150522	155432	3.159	0.424	153449	1.908	0.508
02	020	Anchorage Borough	226338	231238	2.119	0.671	232174	2.514	0.518
04	013	Maricopa County	2122101	2180538	2.680	0.496	2160697	1.786	0.512
04	019	Pima County	666880	686848	2.907	0.486	681920	2.206	0.464
04	021	Pinal County	116379	121955	4.572	0.577	120033	3.045	0.584
04	025	Yavapai County	107714	110720	2.715	0.575	109685	1.797	0.442
04	027	Yuma County	106895	111958	4.522	0.570	109960	2.787	0.572
05	119	Pulaski County	349660	360243	2.938	0.432	357441	2.177	0.517
05	143	Washington County	113409	116428	2.593	0.474	115578	1.877	0.615
06	001	Alameda County	1279182	1323971	3.383	0.455	1317233	2.889	0.5xx
06	007	Butte County	182120	187906	3.079	0.548	186831	2.522	0.554
06	013	Contra Costa County	803732	825024	2.581	0.603	817943	1.737	0.400
06	017	El Dorado County	125995	126797	0.633	0.696	128413	1.883	0.451
06	019	Fresno County	667490	692890	3.666	0.457	691987	3.540	0.501
06	023	Humbolt County	119118	122156	2.487	0.582	122410	2.689	0.488
06	025	Imperial County	109303	116024	5.793	0.705	113220	3.460	0.866
06	029	Kern County	543477	566235	4.019	0.473	558755	2.734	0.375
06	031	Kings County	101469	105597	3.909	0.504	105099	3.454	0.581
06	037	Los Angeles County	8863164	9291955	4.615	0.448	9168889	3.334	0.548
06	041	Marin County	230096	232036	0.836	0.651	232947	1.224	0.523
06	047	Merced County	178403	186707	4.448	0.470	185406	3.777	0.628
06	053	Monterey County	355660	370124	3.908	0.441	367580	3.243	0.644
06	055	Napa County	110765	113411	2.333	0.503	113298	2.236	0.447
06	059	Orange County	2410556	2469336	2.380	0.519	2461373	2.065	0.493
06	061	Placer County	172796	174772	1.131	0.575	175303	1.430	0.374
06	065	Riverside County	1170413	1220764	4.125	0.487	1198964	2.381	0.343
06	067	Sacramento County	1041219	1069918	2.682	0.491	1065198	2.251	0.524
06	071	San Bernardino County	1418380	1490697	4.851	0.501	1455550	2.554	0.355
06	073	San Diego County	2498016	2576888	3.061	0.442	2560392	2.436	0.486
06	075	San Francisco County	723959	756182	4.261	0.504	745573	2.899	0.626

06	077	San Joaquin County	480628	498718	3.627	0.453	495154	2.934	0.381
06	079	San Luis Obispo County	217162	222991	2.614	0.513	222841	2.549	0.500
06	081	San Mateo County	649623	664465	2.234	0.571	561709	1.826	0.457
06	083	Santa Barbara County	369608	383034	3.505	0.473	381039	3.000	0.645
06	085	Santa Clara County	1497577	1544157	3.017	0.453	1531196	2.196	0.475
06	087	Santa Cruz County	229734	238267	3.581	0.503	236007	2.658	0.531
06	089	Shasta County	147036	150573	2.349	0.528	150145	2.070	0.447
06	095	Solano County	340421	353913	3.812	0.450	348512	2.322	0.324
06	097	Sonoma County	388222	399078	2.720	0.504	397377	2.304	0.422
06	099	Stanislaus County	370522	382342	3.092	0.487	380699	2.673	0.475
06	107	Tulane County	311921	324294	3.815	0.482	323520	3.585	0.681
06	111	Ventura County	669016	694637	3.688	0.468	683672	2.144	0.357
06	113	Yolo County	141092	145883	3.284	0.456	145975	3.345	0.565
08	001	Adams County	265038	271716	2.458	0.654	269856	1.786	0.496
08	005	Arapahoe County	391511	398166	1.671	0.693	397542	1.517	0.588
08	013	Boulder County	225339	229447	1.790	0.591	230754	2.347	0.533
08	031	Denver County	467610	482714	3.129	0.579	480862	2.756	0.498
08	041	El Paso County	397014	407843	2.655	0.493	405212	2.023	0.558
08	059	Jefferson County	438430	444327	1.327	0.706	442890	1.007	0.577
08	069	Larimer County	186136	189346	1.695	0.596	190569	2.326	0.527
08	101	Pueblo County	123051	125654	2.072	0.550	125754	2.149	0.540
08	123	Weld County	131821	134887	2.273	0.534	135793	2.925	0.518
09	001	Fairfield County	827645	831105	0.416	0.593	832682	0.605	0.384
09	003	Hartford County	851783	857182	0.630	0.589	857897	0.713	0.483
09	005	Litchfield County	174092	175581	0.848	0.538	175080	0.565	0.523
09	007	Middlesex County	143196	143812	0.428	0.529	143825	0.437	0.526
09	009	New Haven County	804219	807947	0.461	0.583	807987	0.466	0.514
09	011	New London County	254957	255796	0.328	0.554	257535	1.001	0.470
09	013	Tolland County	128699	129683	0.759	0.599	129510	0.626	0.561
09	015	Windham County	102525	104554	1.941	0.823	103793	1.222	0.681
10	001	Kent County	110993	114068	2.696	0.443	112995	1.772	0.394
10	003	New Castle County	441946	456338	3.154	0.510	450294	1.854	0.516
10	005	Sussex County	113229	116255	2.603	0.501	115083	1.611	0.452
11	001	District of Columbia	606900	638747	4.986	0.517	628309	3.407	0.901
12	001	Alachua County	181596	188223	3.521	0.429	186051	2.394	0.635
12	005	Bay County	126994	130912	2.993	0.477	129096	1.629	0.536
12	009	Brevard County	398978	410499	2.807	0.446	404953	1.476	0.445
12	011	Broward County	1255488	1291812	2.812	0.453	1277394	1.715	0.529
12	015	Charlotte County	110975	112871	1.680	0.526	111898	0.825	0.353

12	019	Clay County	105986	106804	0.766	0.595	107762	1.648	0.376
12	021	Collier County	152099	156294	2.684	0.526	154958	1.845	0.464
12	025	Dade County	1937094	1997643	3.031	0.591	2011300	3.690	0.945
12	031	Duval County	672971	697735	3.549	0.463	687821	2.159	0.549
12	033	Escambia County	262798	271007	3.029	0.466	268329	2.061	0.495
12	053	Hernando County	101115	100975	-0.139	0.612	102051	0.918	0.319
12	057	Hillsborough County	834054	851877	2.092	0.448	853411	2.268	0.478
12	069	Lake County	152104	155095	1.929	0.481	154003	1.233	0.341
12	071	Lee County	335113	343538	2.452	0.465	339589	1.318	0.466
12	073	Leon County	192493	199708	3.613	0.437	196621	2.100	0.615
12	081	Manatee County	211707	216819	2.358	0.508	214609	1.352	0.513
12	083	Marion County	194833	199845	2.508	0.487	197743	1.472	0.354
12	085	Martin County	100900	103232	2.259	0.592	102120	1.195	0.406
12	091	Okaloosa County	143776	148410	3.122	0.505	146346	1.756	0.593
12	095	Orange County	677491	700574	3.295	0.458	693622	2.326	0.530
12	097	Osceola County	107728	111188	3.112	0.564	109720	1.816	0.479
12	099	Palm Beach County	863518	886676	2.612	0.484	876764	1.511	0.493
12	101	Pasco County	281131	281049	-0.029	0.614	283694	0.904	0.395
12	103	Pinellas County	851659	861306	1.120	0.448	860438	1.020	0.555
12	105	Polk County	405382	416923	2.768	0.470	411918	1.587	0.405
12	111	St. Lucie County	150171	154362	2.715	0.479	152554	1.562	0.474
12	115	Sarasota County	277776	283554	2.038	0.550	279921	0.766	0.505
12	117	Seminole County	287529	297007	3.191	0.569	292736	1.779	0.505
12	127	Volusia County	370712	380601	2.598	0.512	375737	1.338	0.463
13	021	Bibb County	149967	154963	3.224	0.453	157035	2.005	0.475
13	051	Chatham County	216935	224122	3.207	0.435	221102	1.885	0.506
13	063	Clayton County	182052	184137	1.132	0.562	186841	2.563	0.581
13	067	Cobb County	447745	453535	1.277	0.544	456480	1.914	0.547
13	089	DeKalb County	545837	553706	1.421	0.533	561155	2.730	0.608
13	121	Fulton County	648951	671488	3.356	0.442	668695	2.953	0.738
13	135	Gwinnett County	352910	356619	1.040	0.611	359473	1.826	0.488
13	215	Muscogee County	179278	185474	3.341	0.505	183097	2.086	0.554
13	245	Richmond County	189719	195914	3.162	0.443	194873	2.645	0.584
15	001	Hawaii County	120317	121720	1.153	0.717	122654	1.905	0.750
15	003	Honolulu County	836231	861245	2.904	0.570	852074	1.859	0.837
15	009	Maui County	100374	101591	1.198	0.714	102187	1.774	0.741
16	001	Ada County	205775	208426	1.272	0.594	209575	1.813	0.463
17	019	Champaign County	173025	177031	2.263	0.553	175375	1.340	0.414
17	031	Cook County	5105067	5212195	2.055	0.423	5186429	1.569	0.574

17	043	DuPage County	781666	789453	0.986	0.499	784956	0.419	0.399
17	089	Kane County	317471	324570	2.187	0.524	320253	0.869	0.413
17	097	Lake County	516418	524672	1.573	0.558	519660	0.624	0.330
17	099	LaSalle County	106913	106411	0.472	0.538	107150	0.222	0.416
17	111	McHenry County	183241	184777	0.831	0.510	183780	0.293	0.397
17	113	McLean County	129180	131827	2.008	0.582	130128	0.729	0.408
17	115	Macon County	117206	119550	1.961	0.570	117856	0.551	0.357
17	119	Madison County	249238	251156	0.764	0.432	250446	0.483	0.305
17	143	Peoria County	182827	186534	1.987	0.534	184180	0.735	0.372
17	161	Rock Island County	148723	151424	1.784	0.534	149787	0.711	0.451
17	163	St. Clair County	262852	266701	1.443	0.423	266421	1.340	0.409
17	167	Sangamon County	178386	181578	1.758	0.542	179149	0.426	0.399
17	179	Tazwell County	123692	124872	0.945	0.561	123942	0.202	0.407
17	197	Will County	357313	363530	1.710	0.554	359200	0.525	0.284
17	201	Winnebago County	252913	257702	1.858	0.528	254302	0.546	0.378
18	003	Allen County	300836	306760	1.931	0.534	302274	0.476	0.392
18	035	Delaware County	119659	121730	1.701	0.537	120341	0.566	0.402
18	039	Elkhart County	156198	158664	1.554	0.530	156797	0.382	0.443
18	057	Hamilton County	108936	109674	0.673	0.513	109211	0.252	0.385
18	089	Lake County	475594	487249	2.392	0.552	480322	0.984	0.427
18	091	LaPorte County	107066	107036	-0.028	0.462	107368	0.281	0.480
18	095	Madison County	130669	132535	1.408	0.514	131090	0.321	0.403
18	097	Marion County	797159	803890	0.837	0.577	808143	1.359	0.523
18	127	Porter County	128932	130035	0.848	0.659	129287	0.274	0.397
18	105	Monroe County	108928	111084	1.896	0.552	110094	-1.013	0.498
18	141	St. Joseph County	247052	251786	1.880	0.535	248403	0.544	0.355
18	157	Tippecanoe County	130598	133031	1.829	0.550	132098	1.135	0.459
18	163	Vanderburgh County	165058	168249	1.897	0.596	165711	0.394	0.418
18	167	Vigo County	106107	107712	1.490	0.517	106607	0,469	0.398
19	013	Black Hawk County	123798	126453	2.100	0.553	124529	0.587	0.373
19	113	Linn County	168767	171900	1.823	0.541	169329	0.332	0.387
19	153	Polk County	327140	334027	2.062	0.537	329530	0.725	0.432
19	163	Scott County	150979	154206	2.093	0.533	152246	0.832	0.431
20	091	Johnson County	355054	358386	0.930	0.435	357029	0.553	0.418
20	173	Sedgwick County	403662	409349	1.389	0.407	407780	1.010	0.440
20	177	Shawnee County	160976	164773	2.304	0.525	161845	0.537	0.394
20	209	Wyandotte County	161993	165674	2.222	0.456	164206	1.348	0.460
21	067	Fayette County	225366	233157	3.342	0.602	229930	1.985	0.705
21	111	Jefferson County	664937	685007	2.930	0.439	676776	1.749	0.537

21	117	Kenton County	142031	145523	2.400	0.593	144235	1.528	0.552
22	017	Caddo Parish	248253	256120	3.072	0.428	254356	2.400	0.529
22	019	Calcasieu Parish	168134	172829	2.717	0.405	170974	1.661	0.420
22	033	East Baton Rouge Parish	380105	392277	3.103	0.395	390145	2.574	0.569
22	051	Jefferson Parish	448306	458990	2.326	0.470	457937	2.103	0.525
22	055	Lafayette Parish	164762	169813	2.974	0.409	168125	2.000	0.497
22	071	Orleans Parish	496938	514558	3.424	0.486	513936	3.307	0.876
22	073	Quachita Parish	142191	146297	2.807	0.400	144953	1.905	0.438
22	079	Rapides Parish	131556	135085	2.612	0.389	133995	1.820	0.399
22	103	St. Tammany Parish	144508	147804	2.230	0.451	146874	1.611	0.365
23	001	Androscoggin County	105259	104912	-0.331	0.585	106120	0.812	0.511
23	005	Cumberland County	243135	243615	0.197	0.539	245246	0.861	0.524
23	011	Kennebec County	115904	117501	1.359	0.693	116582	0.581	0.505
23	019	Penobscot County	146601	147574	0.659	0.563	147738	0.770	0.532
23	031	York County	164587	166105	0.914	0.552	165635	0.633	0.520
24	003	Anne Arundel County	427239	431624	1.016	0.537	434447	1.659	0.406
24	005	Baltimore County	692134	696225	0.588	0.567	702812	1.519	0.507
24	013	Carroll County	123372	124098	0.585	0.606	124911	1.232	0.459
24	017	Charles County	101154	102192	1.016	0.571	102794	1.595	0.403
24	021	Frederick County	150208	152604	1.570	0.494	152690	1.626	0.431
24	025	Hartford County	182132	183499	0.745	0.583	185018	1.560	0.359
24	027	Howard County	187328	189033	0.902	0.582	190409	1.618	0.466
24	031	Montgomery County	757027	764514	0.979	0.563	771160	1.833	0.482
24	033	Prince Georges County	729268	740060	1.458	0.579	751587	2.970	0.627
24	043	Washington County	121393	124802	2.732	0.464	123237	1.496	0.460
24	510	Baltimore City	736014	772082	4.672	0.511	759127	3.045	0.808
25	001	Barnstable County	186605	189889	1.729	0.855	187904	0.691	0.530
25	003	Berkshire County	139352	139722	0.265	0.520	140508	0.823	0.505
25	005	Bristol County	506325	505255	-0.212	0.554	509637	0.650	0.452
25	009	Essex County	670080	670474	0.059	0.579	671451	0.204	0.466
25	013	Hampden County	456310	457899	0.347	0.585	458054	0.381	0.706
25	015	Hampshire County	146568	147943	0.929	0.563	147848	0.866	0.555
25	017	Middlesex County	1398468	1402907	0.316	0.600	1399207	0.053	0.615
25	021	Norfolk County	616087	618087	0.324	0.653	611139	-0.810	0.744
25	023	Plymouth County	435276	436386	0.254	0.580	436400	0.258	0.406
25	025	Suffolk County	663906	670095	0.924	0.744	680818	2.484	0.777
25	027	Worcester County	709705	711256	0.218	0.537	713339	0.509	0.456
26	017	Bay County	111723	113132	1.245	0.537	111895	0.153	0.450
26	021	Berrien County	161378	163661	1.395	0.598	162674	0.796	0.454

26	025	Calhoun County	135982	138148	1.568	0.517	136672	0.505	0.398
26	049	Genesse County	430459	438800	1.901	0.538	434600	0.953	0.414
26	065	Ingham County	281912	288505	2.285	0.534	286089	1.460	0.534
26	075	Jackson County	149756	151533	1.173	0.526	150189	0.288	0.511
26	077	Kalamazoo County	223411	227212	1.673	0.520	224957	0.687	0.406
26	081	Kent County	500631	509273	1.697	0.526	504353	0.738	0.407
26	093	Livingston County	115645	116408	0.656	0.511	115499	-0.126	0.949
26	099	Macomb County	717400	722597	0.719	0.522	718766	0.190	0.387
26	115	Monroe County	133600	134642	0.774	0.511	133783	0.137	0.577
26	121	Muskegon County	158983	161494	1.555	0.535	159784	0.501	0.394
26	125	Oakland County	1083592	1094932	1.036	0.481	1088374	0.439	0.383
26	139	Ottawa County	187768	189955	1.151	0.605	188460	0.367	0.443
26	145	Saginaw County	211946	216155	1.947	0.537	213567	0.759	0.401
26	147	St. Clair County	145607	147341	1.177	0.440	145854	0.169	0.512
26	161	Washtenaw County	282937	288679	1.989	0.516	286038	1.084	0.427
26	163	Wayne County	2111687	2160354	2.253	0.426	2144482	1.529	0.478
27	003	Anoka County	243641	245862	0.903	0.517	244251	0.250	0.375
27	037	Dakota County	275227	278038	1.011	0.512	276471	0.450	0.389
27	053	Hennepin County	1032431	1044852	1.189	0.381	1041265	0.848	0.467
27	109	Olmsted County	106470	108411	1.790	0.553	106753	0.265	0.411
27	123	Ramsey County	485765	491319	1.130	0.382	490387	0.943	0.479
27	137	St. Louis County	198213	201605	1.683	0.576	198462	0.126	0.430
27	145	Stearns County	118791	121193	1.982	0.639	119274	0.405	0.560
27	163	Washington County	145896	147156	0.856	0.506	146053	0.108	0.344
28	047	Harrison County	165365	170273	2.882	0.422	168426	1.818	0.509
28	049	Hinds County	254441	264818	3.919	0.446	261731	2.785	0.609
28	059	Jackson County	115243	118271	2.560	0.460	117089	1.576	0.407
29	019	Boone County	112379	115311	2.543	0.550	113620	1.092	0.444
29	047	Clay County	153411	154746	0.863	0.396	154298	0.575	0.414
29	077	Greene County	207949	211970	1.897	0.545	208941	0.475	0.429
29	095	Jackson County	633232	645060	1.834	0.378	640624	1.154	0.466
29	099	Jefferson County	171380	172865	0.859	0.510	171632	0.147	0.504
29	183	St. Charles County	212907	215015	0.980	0.431	213851	0.442	0.380
29	189	St. Louis County	993529	1010023	1.633	0.458	999753	0.623	0.370
29	510	St. Louis City	396685	408263	2.836	0.518	405175	2.096	0.682
30	111	Yellowstone County	113419	114710	1.125	0.605	115539	1.835	0.450
31	055	Douglas County	416444	421918	1.297	0.419	420353	0.930	0.453
31	109	Lancaster County	213641	218226	2.101	0.611	215022	0.642	0.420
31	153	Sarpy County	102583	104050	1.410	0.492	103780	1.154	0.483

32	003	Clark County	741459	759866	2.422	0.518	758692	2.271	0.521
32	031	Washoe County	254667	258898	1.634	0.556	261007	2.429	0.510
33	011	Hillsborough County	336073	335652	-0.125	0.578	338911	0.838	0.500
33	013	Merrimack County	120005	121598	1.310	0.636	120910	0.748	0.539
33	015	Rockingham County	245845	246967	0.454	0.586	247556	0.691	0.546
33	017	Strafford County	104233	104021	-0.204	0.583	105081	0.807	0.557
34	001	Atlantic County	224327	227837	1.541	0.546	226943	1.153	0.374
34	003	Bergen County	825380	829281	0.470	0.580	820928	-0.542	0.786
34	005	Burlington County	395066	401239	1.539	0.665	394939	-0.032	0.568
34	007	Camden County	302824	510058	1.418	0.621	503429	0.120	0.719
34	011	Cumberland County	138053	140210	1.538	0.530	139656	1.148	0.379
34	013	Essex County	778206	802268	2.999	0.560	799678	2.685	0.782
34	015	Glouster County	230082	233020	1.261	0.699	229106	-0.426	0.624
34	017	Hudson County	553099	568477	2.705	0.577	569258	2.839	1.107
34	019	Hunterdon County	107776	107861	0.079	0.603	108451	0.623	0.745
34	021	Mercer County	325824	331440	1.694	0.544	328647	0.859	0.554
34	023	Middlesex County	671780	677682	0.871	0.548	672992	0.180	0.712
34	025	Monmouth County	553124	556412	0.591	0.574	550805	-0.421	0.687
34	027	Morris County	421353	425501	0.975	0.717	419138	-0.529	0.670
34	029	Ocean County	433203	433516	0.072	0.599	429899	-0.769	0.702
34	031	Passaic County	453060	461845	1.902	0.541	459194	1.336	0.858
34	035	Somerset County	240279	241669	0.575	0.578	239512	-0.320	0.617
34	037	Sussex County	130943	132073	0.856	0.729	131218	0.210	0.539
34	039	Union County	493819	503004	1.826	0.588	497433	0.727	0.778
35	001	Bernalillo County	480577	497633	3.427	0.518	491854	2.293	0.457
35	013	Dona Ana County	135510	141574	4.283	0.545	139939	3.165	0.665
36	001	Albany County	292594	295111	0.853	0.530	293849	0.427	0.656
36	005	Bronx County	1203789	1245874	3.378	0.730	1265768	4.897	1.410
36	007	Broome County	212160	212548	0.183	0.541	213689	0.716	0.458
36	013	Chautauqua County	141895	141997	0.072	0.525	143047	0.805	0.539
36	027	Dutchess County	259462	261192	0.662	0.543	261808	0.896	0.459
36	029	Erie County	968532	976594	0.826	0.588	969213	0.070	0.650
36	045	Jefferson County	110943	112132	1.060	0.562	112635	1.503	0.718
36	047	Kings County	2300664	2379894	3.329	0.592	2389150	3.704	0.906
36	055	Monroe County	713968	722929	1.240	0.536	716126	0.301	0.641
36	059	Nassau County	1287348	1296128	0.677	0.571	1277449	-0.775	0.827
36	061	New York County	1487536	1537991	3.281	0.596	1541441	3.497	0.969
36	063	Niagra County	220756	221792	0.467	0.537	220729	-0.012	0.512
36	065	Oneida County	250836	251805	0.385	0.510	252906	0.819	0.447

36	067	Onondaga County	468973	472839	0.818	0.532	469750	0.165	0.638
36	071	Orange County	307647	309752	0.680	0.564	310882	1.040	0.451
36	075	Oswego County	121771	121870	0.081	0.623	122882	0.904	0.685
36	081	Queens County	1951598	2004192	2.624	0.624	1992006	2.029	0.806
36	083	Rensselaer County	154429	154995	0.365	0.535	155072	0.415	0.591
36	085	Richmond County	378977	384245	1.371	0.533	378782	-0.052	0.722
36	087	Rockland County	265475	269627	1.540	0.688	264771	-0.266	0.734
36	089	St. Lawrence County	111974	112733	0.673	0.594	113179	1.064	0.684
36	091	Saratoga County	181276	181488	0.117	0.615	181850	0.316	0.500
36	093	Schenectady County	149285	149852	0.378	0.524	148589	-0.468	0.720
36	103	Suffolk County	1321864	1330743	0.667	0.576	1313346	-0.649	0.727
36	111	Ulster County	165304	167147	1.103	0.612	167385	1.244	0.736
36	119	Westchester County	874866	890648	1.772	0.641	879705	0.550	0.687
37	001	Alamance County	108213	111418	2.877	0.439	109811	1.455	0.408
37	021	Buncombe County	174821	179768	2.752	0.465	177162	1.321	0.413
37	035	Catawba County	118412	112063	2.991	0.498	120094	1.401	0.426
37	051	Cumberland County	274566	284189	3,386	0.419	280604	2.152	0.514
37	057	Davidson County	126677	130509	2.936	0.580	128544	1.453	0.455
37	063	Durham County	181835	188378	3.473	0.462	185785	2.126	0.579
37	067	Forsyth County	265878	274462	3.128	0.430	270363	1.659	0.469
37	071	Gaston County	175093	177824	1.536	0.464	177837	1.543	0.456
37	081	Guilford County	347420	358847	3.184	0.443	353615	1.752	0.501
37	119	Mecklenburg County	511433	528981	3.317	0.424	523306	2.269	0.557
37	129	New Hanover County	120284	124111	3.084	0.438	122381	1.714	0.540
37	133	Onslow County	149838	154392	2.950	0.374	153141	2.157	0.415
37	147	Pitt County	107924	110732	2.536	0.423	110516	2.345	0.557
37	151	Randolph County	106546	109790	2.955	0.595	108009	1.354	0.431
37	155	Robeson County	105179	108097	2.699	0.452	107475	2.136	0.534
37	159	Rowan County	110605	111420	0.732	0.524	112305	1.514	0.375
37	183	Wake County	423380	438428	3.432	0.434	432630	2.138	0.493
37	191	Wayne County	104666	107153	2.321	0.401	106769	1.969	0.390
38	017	Cass County	102874	105012	2.036	0.571	103452	0.559	0.461
39	003	Allen County	109755	111410	1.486	0.510	110262	0.460	0.411
39	017	Butler County	291479	295537	1.373	0.535	292902	0.486	0.359
39	023	Clark County	147548	149800	1.503	0.519	148179	0.426	0.406
39	025	Clermont County	150187	151277	0.721	0.514	150784	0.396	0.522
39	029	Columbia County	108276	107516	-0.679	0.584	108375	0.091	0.584
39	035	Cuyahoga County	1412140	1429431	1.210	0.431	1427932	1.106	0.471
39	045	Fairfield County	103461	103995	0.514	0.427	103594	0.129	0.522

39	049	Franklin County	961437	970249	0.908	0.463	975539	1.446	0.539
39	057	Greene County	136731	138166	1.039	0.632	137700	0.704	0.328
39	061	Hamilton County	866228	876347	1.155	0.424	876795	1.205	0.485
39	085	Lake County	215499	216985	0.685	0.519	216122	0.288	0.378
39	089	Licking County	128300	129042	0.575	0.432	128558	0.201	0.519
39	093	Lorain County	271126	275982	1.760	0.520	272668	0.565	0.364
39	095	Lucas County	462361	465553	0.686	0.477	467096	1.014	0.437
39	099	Mahoning County	264806	268995	1.557	0.528	266443	0.614	0.379
39	103	Medina County	122354	123157	0.652	0.514	122484	0.106	0.462
39	113	Montgomery County	573809	583903	1.729	0.528	580267	1.113	0.461
39	133	Portage County	142585	144241	1.148	0.573	143615	0.717	0.542
39	139	Richland County	126137	127829	1.324	0.520	126535	0.314	0.418
39	151	Stark County	367585	372544	1.331	0.525	368829	0.337	0.384
39	153	Summit County	514990	523958	1.712	0.520	518979	0.769	0.415
39	155	Trumbull County	227813	230339	1.097	0.560	228736	0.403	0.397
39	165	Warren County	113909	114657	0.652	0.498	114158	0.218	0.364
39	169	Wayne County	101461	100828	-0.628	0.605	101745	0.279	0.620
39	173	Wood County	113269	113881	0.537	0.446	113912	0.565	0.418
40	027	Cleveland County	174253	178292	2.265	0.466	177845	2.020	0.539
40	031	Comanche County	111486	114833	2.915	0.418	113756	1.996	0.506
40	109	Oklahoma County	599611	613697	2.295	0.419	612788	2.150	0.547
40	143	Tulsa County	503341	514637	2.195	0.453	512955	1.874	0.534
41	005	Clackamas County	278850	279977	0.403	0.724	281892	1.079	0.452
41	029	Jackson County	146389	150125	2.489	0.537	149287	1.941	0.441
41	039	Lane County	282912	289415	2.247	0.551	289266	2.197	0.493
41	047	Marion County	228483	234494	2.563	0.508	233587	2.185	0.434
41	051	Multnomah County	583887	598049	2.368	0.489	593788	1.668	0.652
41	067	Washington County	311554	314044	0.793	0.688	315806	1.346	0.623
42	003	Allegheny County	1336449	1346520	0.748	0.600	1331707	-0.356	0.758
42	007	Beaver County	186093	186376	0.152	0.593	185256	-0.452	0.637
42	011	Berks County	336523	337434	0.270	0.536	338569	0.604	0.426
42	013	Blair County	130542	130430	-0.086	0.532	131077	0.408	0.448
42	017	Bucks County	541174	545735	0.836	0.726	537873	-0.614	0.634
42	019	Butler County	152013	153223	0.790	0.660	152898	0.579	0.635
42	021	Cambria County	163029	162949	-0.049	0.556	163876	0.517	0.481
42	027	Centre County	123786	124397	0.491	0.570	125635	1.472	0.733
42	029	Chester County	376396	380542	1.090	0.704	377088	0.184	0.535
42	041	Cumberland County	195257	195365	0.055	0.575	195256	-0.001	0.547
42	043	Dauphin County	237813	241035	1.337	0.552	239154	0.561	0.577

42	045	Delaware County	547651	554003	1.147	0.694	545064	-0.475	0.771
42	049	Erie County	275572	276888	0.475	0.529	277235	0.600	0.428
42	051	Fayette County	145351	145958	0.416	0.742	146681	0.907	0.808
42	055	Franklin County	121082	122079	0.817	0.632	122180	0.899	0.729
42	069	Lackawanna County	219039	218814	-0.103	0.532	217294	-0.803	0.732
42	071	Lancaster County	422822	423976	0.272	0.564	426528	0.869	0.523
42	075	Lebanon County	113744	113779	0.031	0.543	114518	0.676	0.589
42	077	Lehigh County	291130	291961	0.285	0.515	289980	-0.396	0.661
42	079	Luzerne County	328149	327768	-0.116	0.546	326439	-0.524	0.593
42	081	Lycoming County	118710	118822	0.094	0.538	119511	0.670	0.493
42	085	Mercer County	121003	121190	0.154	0.552	121627	0.513	0.486
42	091	Montgomery County	678111	683019	0.719	0.697	673620	-0.667	0.671
42	095	Northampton County	247105	247686	0.235	0.527	246917	-0.076	0.572
42	101	Philadelphia County	1585577	1606249	1.287	0.609	1608942	1.452	0.742
42	107	Schuylkill County	152585	153416	0.542	0.631	152989	0.264	0.525
42	125	Washington County	204584	205463	0.428	0.738	204548	-0.018	0.506
42	129	Westmoreland County	370321	371539	0.328	0.750	369009	-0.356	0.551
42	133	York County	339574	340569	0.292	0.572	341321	0.512	0.472
44	003	Kent County	161135	161498	0.225	0.654	159355	-1.117	0.776
44	007	Providence County	596270	597016	0.125	0.580	597960	0.283	0.697
44	009	Washington County	110006	110452	0.404	0.638	110982	0.880	0.633
45	003	Aiken County	120940	124770	3.070	0.542	123291	1.907	0.403
45	007	Anderson County	145196	149574	2.927	0.502	147268	1.407	0.373
45	015	Berkley County	128776	133468	3.515	0.555	132081	2.502	0.472
45	019	Charleston County	295039	304829	3.212	0.437	302751	2.547	0.580
45	041	Florence County	114344	118062	3.149	0.453	116745	2.056	0.454
45	045	Greenville County	320167	330290	3.065	0.494	325537	1.650	0.467
45	051	Horry County	144053	147841	2.562	0.452	146650	1.771	0.455
45	063	Lexington County	167611	173083	3.162	0.583	170341	1.602	0.375
45	079	Richland County	285720	295225	3.220	0.421	293299	2.584	0.564
45	083	Spartanburg County	226800	233790	2.990	0.489	230614	1.654	0.374
45	085	Sumter County	102637	105121	2.363	0.403	105017	2.267	0.500
45	091	York County	131497	133960	1.839	0.454	133717	1.660	0.409
46	099	Minnehaha County	123809	126103	1.819	0.578	124220	0.331	0.442
47	037	Davidson County	510784	532433	4.066	0.521	522044	2.157	0.617
47	065	Hamilton County	285536	293917	2.852	0.442	290664	1.764	0.512
47	093	Knox County	335749	345081	2.704	0.466	341481	1.679	0.502
47	125	Montgomery County	100498	104034	3.399	0.463	102468	1.923	0.518
47	149	Rutherford County	118570	122462	3.178	0.466	120716	1.778	0.511

47	157	Shelby County	826330	861616	4.095	0.432	847848	2.538	0.589
47	163	Sullivan County	143596	146794	2.179	0.489	145270	1.152	0.437
47	165	Sumner County	103281	105733	2.319	0.586	104756	1.408	0.343
48	027	Bell County	191088	197377	3.186	0.387	195808	2.410	0.563
48	029	Bexar County	1185394	1220995	2.916	0.498	1230141	3.638	0.744
48	039	Brazoria County	191707	196965	2.670	0.484	195577	1.979	0.374
48	041	Brazos County	121862	126396	3.587	0.520	125880	3.192	0.903
48	061	Cameron County	260120	269903	3.625	0.754	268659	3.178	0.983
48	085	Collin County	264036	271624	2.794	0.479	269149	1.900	0.412
48	113	Dallas County	1852810	1929504	3.975	0.408	1912100	3.101	0.620
48	121	Denton County	273525	282791	3.277	0.444	279483	2.132	0.495
48	135	Ector County	118934	122783	3.135	0.461	121298	1.949	0.583
48	141	El Paso County	591610	611278	3.218	0.611	617397	4.177	0.898
48	157	Fort Bend County	225421	233251	3.357	0.459	230752	2.310	0.338
48	167	Galveston County	217399	223599	2.773	0.388	221787	1.979	0.488
48	183	Gregg County	104948	107799	2.645	0.417	106936	1.860	0.522
48	201	Harris County	2818199	2939388	4.123	0.421	2915587	3.340	0.634
48	215	Hidalgo County	383545	399356	3.959	0.883	399991	4.112	0.841
48	245	Jefferson County	239397	246592	2.918	0.408	243776	1.796	0.441
48	303	Lubbock County	222636	229852	3.139	0.466	228182	2.430	0.599
48	309	McLennan County	189123	194533	2.781	0.393	193347	2.185	0.541
48	329	Midland County	106611	109988	3.070	0.466	108645	1.872	0.498
48	339	Montgomery County	182201	186761	2.442	0.500	185687	1.877	0.441
48	355	Nueces County	291145	299681	2.848	0.533	301959	3.581	0.714
48	423	Smith County	151309	155316	2.580	0.390	154321	1.952	0.391
48	439	Tarrant County	1170103	1212831	3.523	0.405	1200703	2.549	0.540
48	441	Taylor County	119655	123143	2.833	0.479	122112	2.012	0.577
48	453	Travis County	576407	594107	2.979	0.447	596444	3.360	0.663
48	479	Webb County	133239	138180	3.576	0.771	137203	2.889	1.239
48	485	Wichita County	122378	125621	2.582	0.440	124508	1.711	0.552
48	491	Williamson County	139551	143640	2.847	0.503	142663	2.182	0.376
49	011	Davis County	187941	190520	1.354	0.734	190068	1.119	0.708
49	035	Salt Lake County	725956	736793	1.471	0.635	735135	1.249	0.689
49	049	Utah County	263590	268891	1.971	0.628	271102	2.771	0.691
49	057	Weber County	158330	160566	1.393	0.581	160318	1.240	0.573
50	007	Chittenden County	131761	132031	0.205	0.587	132975	0.913	0.564
51	013	Arlington County	170936	178147	4.048	0.491	175566	2.637	0.724
51	041	Chesterfield County	209274	216590	3.378	0.584	212658	1.591	0.432
51	059	Fairfax County	818584	826402	0.946	0.575	833668	1.809	0.501

51	087	Henrico County	217881	224759	3.060	0.546	221878	1.801	0.506
51	153	Prince William	215686	218414	1.249	0.585	220359	2.121	0.425
51	510	Alexandria City	111183	112748	1.388	0.541	114451	2.856	0.771
51	550	Chesapeake City	151976	153512	1.001	0.556	155185	2.068	0.509
51	650	Hampton City	133793	139284	3.942	0.459	137415	2.636	0.617
51	700	Newport News City	170045	178053	4.498	0.468	175121	2.899	0.689
51	710	Norfolk City	261229	273457	4.472	0.444	269011	2.893	0.733
51	740	Portsmouth City	103907	108477	4.213	0.474	106837	2.742	0.695
51	760	Richmond City	203056	209959	3.288	0.549	208987	2.838	0.817
51	810	Virginia Beach City	393069	408213	3.710	0.487	402092	2.244	0.558
53	005	Benton County	112560	115161	2.259	0.556	115073	2.184	0.445
53	011	Clark County	238053	245741	3.129	0.555	241186	1.299	0.533
53	033	King County	1507319	1536441	1.895	0.519	1531673	1.590	0.612
53	035	Kitsap County	189731	196029	3.213	0.531	193702	2.050	0.425
53	053	Pierce County	586203	607187	3.456	0.502	597344	1.865	0.541
53	061	Snohomish County	465642	470715	1.078	0.625	471683	1.281	0.537
53	063	Spokane County	361364	370081	2.355	0.539	365976	1.260	0.577
53	067	Thurston County	161238	166421	3.114	0.542	164464	1.962	0.425
53	073	Whatcom County	127780	131437	2.782	0.532	130903	2.386	0.487
53	077	Takima County	188823	196444	3.880	0.499	195170	3.252	0.557
54	039	Kanawha County	207619	213488	2.749	0.492	210468	1.354	0.443
55	009	Brown County	194594	197594	1.518	0.540	195417	0.421	0.428
55	025	Dane County	367085	373810	1.799	0.541	370065	0.805	0.441
55	059	Kenosha County	128181	130580	1.837	0.548	128869	0.534	0.392
55	073	Marathon County	115400	116699	1.113	0.555	115646	0.213	0.516
55	079	Milwaukee County	959275	969329	1.037	0.459	975296	1.643	0.590
55	087	Outagamie County	140510	142519	1.410	0.543	141059	0.390	0.428
55	101	Racine County	175034	178398	1.886	0.522	176209	0.667	0.366
55	105	Rock County	139510	141935	1.709	0.558	140129	0.441	0.395
55	117	Sheboygan County	103877	105288	1.340	0.537	104218	0.327	0.445
55	133	Waukesha County	304715	306312	0.521	0.454	305387	0.220	0.361
55	139	Winnebago County	140320	142464	1.505	0.549	140855	0.380	0.418

UC RT Undercount Rate as estimated from the PES.

SE(UCRt) The sampling error of the estimated undercount rate.

January 12, 2000

Chronology of Events for the 1990 Census Adjustment Decision

1980

- 1980 Census taken
- Formation of Undercount Steering Committee
- Decision was made not to adjust the Census. Undercount Research Staff formed to conduct coverage measurement research

1985

- The Bicentennial Census, a National Academy of Sciences Committee on National Statistics report, recommends method for adjustment of the count and improving accuracy in 1990
- Census Bureau tests computer matching technique in Florida test

1986

- Census Bureau conducts Test of Adjustment Related Operations (TARO) in Los Angeles to determine feasibility of adjustment of 1990 census. Bureau concludes that it is technically feasible to adjust the Census, but there are operational concerns.

1987

Spring
- Census Bureau announces it has developed a feasible method for undercount adjustment using a 300,000 household Post Enumeration Survey (PES)

October
- Commerce Department cancels the plans for the PES for adjustment

1988

Fall
- New York City sues the Commerce Dept for reinstatement of the PES and adjustment

1989

July
- Government signs stipulation agreement with New York plaintiffs reinstating PES and adjustment methodology; decision on which results to use set for July 1991. The settlement had 3 outcomes:

    1. A PES would be conducted and evaluated to see if it could be used to correct the Census count by 7/15/91. The Secretary of Commerce would make decision.

    2. The Secretary of Commerce would publish guidelines he would follow to make decision.

    3. An eight member Special Advisory Panel was appointed to advise the Secretary. Four members on each side of the argument.

July	r	Reestablish Census Undercount Steering Committee with Paula Schneider as chair.
1990	r	1990 Census taken
March	r	Guidelines are issued by the Secretary, challenged and upheld
Winter	r	Unadjusted 1990 Census results released
1991		
Jan. 1- April 4	*	Redistricting
June	r	Completed extensive evaluation of the PES. Twenty-one evaluations on the sampling and non-sampling error and eleven for demographic analysis.
June 21	r	Census Undercount Steering Committee issued recommendation that Census be adjusted.
June 28	r	Census Director Barbara Bryant recommended to Secretary Mosbacher that the 1990 Census be adjusted. Special Advisory Panel rendered a split decision (4 members for adjustment and 4 against adjustment).
	r	Under Secretary Darby recommended not to adjust the 1990 Census.
July 22	r	Secretary Mosbacher decided not to adjust the 1990 Census. He directed Bureau to review results of PES and see if PES could be used to adjust the post-censal estimates.
Summer	r	NYC requests a trial to determine if Secretary's decision was "arbitrary and capricious"
August	*	CAPE committee formed to direct research on potential adjustment of post-censal estimates
Fall	*	CAPE issues report indicating that states are improved while results are not conclusive for small areas
	r	Decennial Census Improvement Act creates National Academy of Sciences panel to study improved methods for 2000 Census
October	r	Computer error discovered in 1991 estimates
1992	r	Two CNSTAT panels were formed to study improvements for 2000 Census
Spring	r	1990 Census lawsuit goes to 13 day trial
Dec. 29	r	Census Director Barbara Bryant decides not to adjust post-censal estimates on the basis of the PES. Adjustment would make distribution for states better but cannot make any conclusions for entities with less than 100,000 population. So Director decided not to adjust since could not improve coverage in all areas. She did announce that the Federal Statistical System could use adjusted numbers for survey controls.
1993		
Spring	r	Judge McLaughlin rules, holding that the decision of the Commerce Secretary was not "arbitrary and capricious"

Sept. 29 ▪ BLS decided to use adjusted population counts for CPS controls. Rest of Federal Statistical System followed.

UNITED STATES DEPARTMENT OF COMMER
Bureau of the Census
Washington, D.C. 20233

NOV 25 1992

MEMORANDUM FOR CAPE Committee

From: John H. Thompson
Chief, Decennial Statistical Studies Division

Subject: Addendum to August 7, 1992 CAPE Report

Attached to this memorandum is an addendum to the August 7, 1992 CAPE report. The addendum documents the work that has transpired since the August report was issued.

Attachment

ADDITIONAL RESEARCH ON ACCURACY OF ADJUSTED VERSUS UNADJUSTED 1990 CENSUS BASE FOR USE IN INTERCENSAL ESTIMATES

ADDENDUM TO REPORT OF THE COMMITTEE ON ADJUSTMENT OF POSTCENSAL ESTIMATES BUREAU OF THE CENSUS DEPARTMENT OF COMMERCE NOVEMBER 25, 1992

The purpose of this addendum is to summarize and document additional research conducted to examine the accuracy of a potential adjustment to the 1990 census base for use in producing intercensal estimates. The August 7, 1992 report of the CAPE, and the subsequent discussions documented in the meeting notes describe a number of areas where the committee felt more information would be helpful to the decision process. The decision to extend the period of outside comment has enabled some additional research to be carried out to more thoroughly explore a subset of these questions. This addendum summarizes that additional research.

The research has been in three basic areas -- additional analysis of accuracy based on loss functions, additional study of homogeneity within post strata, and additional work based on demographic analysis.

1. ADDITIONAL RESEARCH BASED ON LOSS FUNCTION ANALYSIS

The additional research on loss function analysis has fallen into three basic areas. First, loss functions were computed to study the accuracy of the 27 cities larger than the smallest state, Wyoming. Secondly, loss functions were computed to compare additional distributions of population shares. Finally, we computed loss functions to study some of the properties of a composite estimator suggested at the September 4, 1992 CAPE meeting.

It must be noted that the committee had significant concerns regarding the construction of the target populations which serve as the standard in loss function analysis to assess the accuracy of the adjusted and unadjusted census data. The research described below has not addressed this concern.

1.1. Cities Larger Than Wyoming

A loss function analysis was conducted to study the accuracy of adjustment for cities larger than the state of Wyoming.[1] The analysis was conducted to study the accuracy of an adjustment on the distribution of proportionate shares for just the 27 cities, for the distribution of shares for the 27 cities and the balance of the United States, and a state-by-state study of the within state population shares of the cities within the state larger than Wyoming and the state balances. This research was motivated to assess the accuracy of a suggested adjustment process that would have adjusted states and only cities larger in size than the smallest state. The following results were observed:

> The hypothesis tests at the 10-percent significance level did not indicate an improvement from adjustment to the distribution of the population shares among the 27 cities. The committee discussed this result and noted that most of these cities had high undercounts relative to the national average, but there was not a large degree of difference in the undercount among the cities.
>
> The hypothesis tests indicated that the distribution of the shares for the 27 cities and the balance of the United States was improved. Here, the committee noted that there was a difference between the undercount for the large cities, and the combined balance of the United States.
>
> The state-by-state comparisons of the within state distribution of population shares for the large cities and the corresponding state balances was mixed. For example, for the states of New York and Massachusetts the hypothesis tests indicated an improvement in the population shares between the big city and the balance of the state. These test results were observed for each of the methods of computing the target values and each of the methods of computing the loss functions. For the remaining states with one of the 27 large cities, significance at the 10 percent level was not consistently observed for each method.

---

[1]This work is documented in detail in the CAPE minutes 9-1-92, Attachment 1. Details of loss function analysis appear in "Loss Function Analysis for the Post Census Review (PCR) Estimates," Mary Mulry, 7-2-92. Cities larger than Wyoming were selected because of concerns about only adjusting states when these cities had comparable reliability.

1.2. Additional Distributions of Population Shares[2]

One criticism of the loss function analysis has been that we had been restricting our examination to the distribution of population shares for entities within specific size categories (e.g., places with population of 100,000 or more) rather than computing the loss function analysis on a distribution which includes all places or counties. We have addressed this concern by computing loss functions and associated hypothesis tests for three additional distributions of population shares:

(1) All counties;

(2) All places with 100,000 or more population and the 50 state balances of areas not included in a place with 100,000 or more population;

(3) All places with 25,000 or more population, and the balances of counties not in a place with 25,000 or more population.

Each of these three distributions completely partition the entire population of the United States. These results are discussed in section 1.4, below.

1.3. Raked Composite Estimator

Another criticism was that the composite estimation (option 4)[3] depressed the effect of adjustment among demographic groups at the national level. Therefore, a composite estimation methodology based on controlling the "50-50" estimator to national controls obtained from the Post-Enumeration Survey (PES) was also studied.[4] Eight Race/Hispanic Origin categories crossed with tenure were used as controls:

Non-Hispanic White and Other by Owner and Non-owner (2)

Black by Owner and Non-owner (2)

Non-Black Hispanic by Owner and Non-owner (2)

---

[2] A description of the details of this research is in "Additional Loss Function Analysis," John Thompson, November 4, 1992 Memorandum For CAPE Committee.

[3] This option is discussed in the August 10, 1992 Federal Register.

[4] This estimation technique is described in "Additional Loss Function Analysis," John Thompson, November 4, 1992 Memorandum for CAPE Committee.

Asian and Pacific Islanders (1)

American Indians on Reservations (1)

The controlling was carried out within each of the above categories by first calculating the difference between the full adjustment and the original 50-50 composite estimator at the national level. This difference was then allocated sub-nationally within the control categories using the proportional distribution of the original 50-50 composite estimator. Statistically, this follows a technique referred to as "raking", leading to the terminology of Raked Composite Estimator. These results are discussed below.

1.4. Summary of Results

The results of the loss function analysis for the measures described in sections 1.2 and 1.3 are summarized in the form of significance values for the hypothesis testing. We have restricted our analysis to two of the methods of calculating the target populations -- PROPUC and GROSDSE without correlation bias. (These methods of computing the targets are described in "Total Error for Postcensus Review Estimates of Population" by Mary Mulry, July 7, 1992). We were not able to carry out the analysis for all of the methods of computing the target populations. We selected these two methods for study because we believe that they will cover the range of alternative methods to calculate the target populations. We also excluded the correlation bias modeling because there were still many questions about how to estimate correlation bias and how to adequately allocate the estimate of correlation bias to all geographic areas of interest. The effect of not including correlation will also tend to be conservative, since including measures of correlation bias would most likely favor adjustment.

1.4.1 Weighted Squared Error

This section summarizes the CAPE presentation and discussion of these results. The committee discussion was centered on weighted squared error results. Table 1 presents the significance values for the weighted squared error loss function hypothesis test results.[5] We show data for the full adjustment, the raked composite, and the 50-50 composite. The results are displayed for states, counties, and places. We show the results for the previous size category distributions, and for the three new distributions. A summary of key results follows:

---

[5]These results are discussed in more detail in "Additional Loss Function Analysis," John Thompson, November 4, 1992 Memorandum for CAPE Committee.

(1) The hypothesis test significance values for the full adjustment indicate little evidence of an improvement from adjustment for the PROPUC target population method for most size categories -- they are well above 0.10. However, for the loss functions reflecting places of 100,000 or more population and state balances (the last line in Table 1), the significance level approaches 0.10.

(2) The hypothesis tests for the GROSDSE target population method are much more significant than for the PROPUC method, thus indicating more evidence for improvement due to adjustment. This is particularly the case for the new size categories for counties and places.

(3) The hypothesis tests for the raked composite are much more significant than for the full adjustment, indicating more evidence for improvement. These tests are significant at the 10-percent level for the new size categories, for both target population methodologies.

(4) The significance values for the 50-50 composite (where available) are similar to the raked composite estimator.

(5) The hypothesis tests for places between 50,000 and 99,999 population are more significant than the tests for areas with 100,000 or more population.

1.4.2 Squared Error and Relative Squared Error

Tables 2 and 3 give significance values for loss functions based on squared error and relative squared error, respectively.

1.5. Summary of Committee Discussion

The committee discussed these data and noted that while gains had been achieved in reducing sampling error, the raked composite estimator depended more heavily on the assumption of homogeneity (discussed in more detail below). Many on the committee expressed concern with balancing the reduction in sampling error with the greater dependence on the homogeneity assumption. Given these concerns, there was general consensus that the raked composite estimator offered great potential for future research. However, there was not currently enough information available to select this estimator as superior to the full adjustment which had been more thoroughly studied and discussed.

The committee noted that there was some evidence that large areas (greater than 100,000 population) were improved by adjustment when compared to the balance of state. However, the committee also noted that these loss function results should be treated with caution, since they were subject to the same limitations as noted in the August 7, 1992 committee report.

2. ADDITIONAL RESEARCH ON THE HOMOGENEITY ASSUMPTION

The validity of the homogeneity assumption was one of five basic issues addressed by the CAPE report of August 7, 1992. The report summarized the status of knowledge at that date by the following (p. 25):

> Summary: Just as in July 1991, the results on whether the homogeneity assumption holds are inconclusive. The new research used to examine the homogeneity assumption (called artificial population analysis) indicates that the assumption does not hold when the bias in the estimate gets to be about 25% or higher. Since the bias in the Post Enumeration Survey (PES) estimate as currently measured is 22% to 45%, the Committee was concerned.

New analysis has refined the use of the artificial population analysis to examine quantitatively the effect of departures from the homogeneity assumption and to assess the performance of the loss function analysis in the presence of heterogeneity. Reexamination of the evidence has identified areas of incompleteness in the analysis of the previous findings about the effect of bias on the loss function analysis.

2.1. Refinements to the Analysis of the Artificial Populations

Much of the previous research of the artificial populations focused on assessment of whether, in the absence of sampling variance and bias, the 357-post-strata estimator would represent an improvement in the true distribution over the census distribution. This research left largely unanswered questions about the possible size of the effect that departures from the homogeneity (or synthetic) assumption could have, and how such departures would interact with other aspects of the PES analysis, especially the loss function analysis. The results of the reanalysis described below were presented to the committee on November 5, 1992. The committee has not conducted an extensive discussion of these new findings.

The reanalysis focused on three measures: squared error, weighted squared error, and relative squared error. (Measures based on absolute error appeared to present difficult technical issues and were not considered.) The analysis was restricted to the state level. A first part of the analysis addressed the question:

> Q1: Compared to other errors in the PES estimation, how much effect could departures from the homogeneity assumption have on the errors of the PES estimates?

This question was addressed by reexpressing previous findings for the artificial populations by forming the ratio of losses under the adjustment compared to no adjustment. Thus, a ratio of 0 would indicate that the homogeneity assumption was completely satisfied, 0.20 indicates that the PES estimator could potentially capture 80 percent of the underlying variation in the corresponding artificial population, a ratio of .80 indicates that adjustment would capture only 20 percent of the underlying variation. Although ratios above 1.0 are theoretically possible, none were observed. The results are included in the minutes of the CAPE for November 5, 1992. Although ratios ran a wide gambit, going down as far as a highly favorable .11, 12 of the 24 ratios exceeded .50. Such evidence indicates a strong possibility that the 357 post-strata design may capture only about half of the true state-to-state variation in undercount.

The strong possibility that errors due to heterogeneity could be as large as half the errors in the census now appears consistent with the observation made by one reviewer that the errors due to heterogeneity could be larger than all of the errors in the PES accounted for by the total error model.

Given this potentially high level of error, it became critical to assess how heterogeneity would affect the loss function analysis. (If heterogeneity was found conclusively to be quite small, then it could be successfully argued that heterogeneity could only have a small impact on the validity of the loss function analysis.) On the other hand, the reanalysis was still consistent with earlier findings, namely, that the bias due to heterogeneity does not, by itself, obviate the ability of the adjustment to make improvements on the census.

The second question is therefore:

> Q2: How does heterogeneity affect the rest of the PES analysis? In particular, in the presence of heterogeneity, can the PES loss function analysis still reliably measure the improvement, if any, that adjustment makes?

The artificial populations were also used to assess this second question. Since the loss function analysis compares the PES estimates and the census to target values constructed through the synthetic estimator used in the PES, the artificial populations can be used to ascertain whether comparison to such targets correctly states, understates, or overstates the actual improvements of adjustment, which are determined by comparing the census and adjusted distribution to the true census values.

Largely, the evidence supported the continued use of the loss function analysis as a measure of the net improvement, although with qualifications. In particular:

- For the majority of populations, the loss function analysis was actually conservative, tending to understate the true improvement in distribution by using target populations constructed from the synthetic model, compared to the actual advantage of adjustment over the census when the true state values were used as a standard for comparison.

- For two populations, poverty and mobility, the balance between the loss function analysis and the actual improvement appeared about right, in some cases overstating the advantage of adjustment slightly.

- In one instance, the artificial population based on unemployment, the synthetic model was the least successful, explaining only about 20 percent of the variability at the state level. Furthermore, the loss function analysis was seriously distorted, presenting a seriously misleading measure of the improvement due to adjustment.

Some attempts had been made to assess the interaction of sampling error on the analysis by assigning sampling variance to the post-strata. Thompson, and Alberti, discussed some of these findings in a memorandum to the CAPE.[6] Their findings indicated that sampling variance would raise serious questions against adjustment. However, the results of the analysis depended on how much variance was assigned. Thompson and Alberti did not have direct estimates of variance for the artificial population variables. In place of arbitrary decisions about variance, Fay, in recent work (memorandum of Nov. 18, 1992 to the CAPE Committee) calculated sample estimates for the 357 post-strata for 5 of the 8 artificial populations, based on the PES sample blocks only, with the appropriate survey weights. These results have not been discussed by the committee. Several members of the committee view them as being more supportive of adjustment, but questions still remain regarding how much variance must be assigned.

2.2. Reexamination of Bias with the Artificial Population Analysis

The previous CAPE report asserted that the artificial population analysis had shown that the improvement from adjustment apparently vanished when the PES estimates were subject to biases on the order of 25 percent, as noted in the cited summary. In fact, reexamination of the findings presented to the committee revealed that the results were different from the interpretation given them in the earlier report.

---

[6]"Additional Results for Artificial Populations'" John Thompson, September 2, 1992 Memorandum for CAPE Committee.

Although most of the CAPE analysis focused on distributive accuracy, the statistical analyses leading to the figure of 25 percent bias were all based on numeric accuracy. Initially, it was thought that the method of modeling the bias would have no effect on the loss functions for population shares. Further analysis has indicated that this is not the case, and that the loss function analysis for population shares is more robust to our method of modeling bias. This work was not available to the committee for discussions regarding the failure of the homogeneity assumption. A more detailed analysis of this work combined with alternative methods of modeling bias should be carried out in future studies to learn more about the effect of bias on the loss function analysis for population shares.

3. ADDITIONAL RESEARCH ON THE CONSISTENCY OF PES ESTIMATES OF COVERAGE WITH DEMOGRAPHIC ANALYSIS AND OTHER INDICATORS OF COVERAGE FOR SUB-STATE AREAS

The CAPE discussed the consistency of the PES and demographic analysis estimates in the August 7, 1992 report. At that point, the committee generally felt that the PES estimates met their face validity expectations at the State level with some individual state exceptions. Since August, additional research has been conducted to examine the face validity of the PES estimates for large sub-state areas based on demographic indicators.

Direct demographic estimates of the population under age 10 were produced and compared to the PES estimates. This work was accomplished in 40 states for 132 large counties and state balances (total of 172 individual areas). Additional work was also carried out for proxy measures of coverage at the sub-state level. Measures such as percent minority, percent renter, substitution rates, and poverty rates were used.

These results were briefly discussed at the November 5, 1992 CAPE meeting.[7] A detailed discussion of these results will be documented in a future internal memorandum from the Population Division. The results tend to indicate that there are very general patterns of agreement between the PES and demographic analysis results. There has been no extensive review or discussion of these findings by the committee, therefore, no conclusions can be stated.

---

[7] This discussion does not appear in the notes of the November 5, 1992 meeting. The results were merely mentioned in passing.

4. FINAL SUMMARY

The additional research described above has addressed some of the concerns documented in the initial report of the committee August 7, 1992. The general conclusions from that report remain much the same.

4.1 The August 7, 1992 report indicates that the committee concluded that on average, an adjusted state base would be more accurate than an unadjusted state base for use in intercensal estimates. This is still the case. The research based on loss functions since August 7, 1992 has indicated that additional evidence exists that adjustment will improve the distribution of population shares for large places (100,000 or more population) compared to the balance of state.

4.2 The research on the homogeneity assumption has indicated that the total error model does not include a complete measure of the error due to failure of the homogeneity assumption. The research also indicated that the loss function analysis based on the total error model was somewhat robust to this problem, and could be viewed as a measure of net improvement. The research also indicated that more information should be gathered regarding the effect of measurement biases on homogeneity.

Table 1 Significance Probabilities for Loss Functions Analysis for the Weighted Squared Error Loss Function

	Full Adjustment		Raked Composite		50:50 Composite	
	PROPUC	GRODSE	PROPUC	GRODSE	PROPUC	GRODSE
STATES						
All States	0.29	0.13	0.13	0.04	0.12	0.05
COUNTIES						
All Counties	0.29	0.12	0.09	0.02	NA	NA
Less than 200K	0.89*	0.61	NA	NA	NA	NA
200K or More	0.06*	0.08	0.04	0.02	0.04	0.02
PLACES						
25K or More	0.27	0.20	0.08	0.05	NA	NA
25K+ County Bal	0.27	0.10	0.08	0.04	NA	NA
50K-100K	0.25*	0.33	0.21	0.10	0.14	0.06
100K or More	0.22*	0.52	0.26	0.25	0.25	0.23
100K+ State Bal	0.14	0.03	0.04	0.01	NA	NA

NOTES:

- PROPUC is the PROPUC without correlation bias target except where indicated otherwise
- GRODSE is the GRODSE without correlation bias target
- An '*' indicates that the PROPUC with correlation bias value is given since the PROPUC without correlation bias value is not available

Table 2 Significance Probabilities for Loss Functions Analysis for the Squared Error Loss Function

	Full Adjustment		Raked Composite		50:50 Composite	
	PROPUC	GRODSE	PROPUC	GRODSE	PROPUC	GRODSE
STATES						
All States	0.21	0.08	0.10	0.03	0.09	0.04
COUNTIES						
All Counties	0.09	0.05	0.04	0.02	NA	NA
Less than 200K	0.84*	0.56	NA	NA	NA	NA
200K or More	0.05*	0.09	0.05	0.03	0.04	0.03
PLACES						
25K or More	0.46	0.49	0.31	0.33	NA	NA
25K+ County Bal	0.14	0.13	0.09	0.10	NA	NA
50K-100K	0.24*	0.32	0.20	0.09	0.13	0.06
100K or More	0.45*	0.81	0.57	0.64	0.57	0.60
100K+ State Bal	0.19	0.04	0.06	0.01	NA	NA

NOTES:

- PROPUC is the PROPUC without correlation bias target except where indicated otherwise
- GRODSE is the GRODSE without correlation bias target
- An '*' indicates that the PROPUC with correlation bias value is given since the PROPUC without correlation bias value is not available

Table 3 Significance Probabilities for Loss Functions Analysis for the Relative Squared Error Loss Function

	Full Adjustment		Raked Composite		50:50 Composite	
	PROPUC	GRODSE	PROPUC	GRODSE	PROPUC	GRODSE
STATES						
All States	0.55	0.26	0.27	0.11	0.26	0.11
COUNTIES						
All Counties	0.90	0.56	0.71	0.33	NA	NA
Less than 200K	0.93*	0.63	NA	NA	NA	NA
200K or More	0.09*	0.10	0.05	0.02	0.06	0.03
PLACES						
25K or More	0.21	0.09	0.05	0.01	NA	NA
25K+ County Bal	0.83	0.57	0.60	0.35	NA	NA
50K-100K	0.25*	0.35	0.22	0.11	0.14	0.07
100K or More	0.13*	0.26	0.15	0.09	0.13	0.09
100K+ State Bal	0.19	0.10	0.06	0.02	NA	NA

NOTES:

- PROPUC is the PROPUC without correlation bias target except where indicated otherwise
- GRODSE is the GRODSE without correlation bias target
- An '*' indicates that the PROPUC with correlation bias value is given since the PROPUC without correlation bias value is not available

11/ 24/92

COMMITTEE ON ADJUSTMENT OF POSTCENSAL ESTIMATES (CAPE)
DISTRIBUTION LIST

B. Bryant	DIR	2049-3
H. Scarr	DIR	2049-3
P. Bounpane	DIR	3049-3
C. Jones	DIR	2031-3
R. Tortora	DIR	2031-3
W. Butz	DIR	2061-3
R. Fay	DIR	3067-3
P. Schneider	FLD	3023-3
D. Blass	FLD	3027-3
M. Mulry	2KS	3525-3
J. Thompson	DSSD	3785-3
J. Long	POP	2019-3
G. Robinson	POP	2022-3
M. Batutis	POP	2336-3
H. Hogan	SRD	3207-4
D. Whitford	DMD	3554-3
T. Walsh	DMD	3586-3
C. Miller	PPDO	2430-3
L. Lyons	PPDO	1085-3

Please direct any changes to Lana Sylier 763-5337.

11/ 24/92

COMMITTEE ON ADJUSTMENT OF POSTCENSAL ESTIMATES (CAPE)
DISTRIBUTION LIST

B. Bryant	DIR	2049-3
H. Scarr	DIR	2049-3
P. Bounpane	DIR	3049-3
C. Jones	DIR	2031-3
R. Tortora	DIR	2031-3
W. Butz	DIR	2061-3
R. Fay	DIR	3067-3
P. Schneider	FLD	3023-3
D. Blass	FLD	3027-3
M. Mulry	2KS	3525-3
J. Thompson	DSSD	3785-3
J. Long	POP	2019-3
G. Robinson	POP	2022-3
M. Batutis	POP	2336-3
H. Hogan	SRD	3207-4
D. Whitford	DMD	3554-3
T. Walsh	DMD	3586-3
C. Miller	PPDO	2430-3
L. Lyons	PPDO	1085-3

Please direct any changes to Lana Sylier 763-5337.

UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, D.C. 20233

NOV 24 199_

MEMORANDUM FOR CAPE Committee

From: John H. Thompson
Chief, Decennial Statistical Studies Division

Subject: Appendix A -- Discussion of Technical Comments

Attached to this memorandum is Appendix A -- discussion of technical comments received from outside reviewers.

Attachment

November 24, 1992

Appendix A: Discussion of Technical Issues Raised by Outside Comment

A number of important technical issues have been raised from the public commentary. Most of these issues have been included in the CAPE discussions and documented in the report of the committee or in the meeting notes.

The majority of the technical comments raised issues against adjustment of the 1990 census base. The concerns expressed were for the most part related to the analysis conducted by the Census Bureau, and the assumptions that underlie the analysis, and the PES estimation. The opinions expressed in support of adjustment generally recommended that the full adjustment be carried out for all levels of geography. A more detailed discussion of these issues is given below:

1. Homogeneity

Several writers pointed out the critical nature of the dependence of the adjustment on the homogeneity assumption. This assumption states that the undercount rates should remain fixed within each of the 357 post-strata. Although post-strata typically fall in several states and numerous counties, within any one post-stratum, undercount rates are assumed to remain fixed from one state to another, from one county to another, etc. The undercount rate is assumed to vary from one post-stratum to another, however; indeed, this variation is the basis of the adjustment. To the extent that the homogeneity assumption is violated, the undercount rates are said to be "heterogeneous." By adjusting all persons in a post-stratum by the same factor, the PES estimator assumes that the homogeneity assumption holds.

When homogeneity is defined in this manner, it is virtually self-evident that the assumption cannot hold exactly. The key issue, from the perspective of CAPE, is whether the homogeneity assumption represents an adequate approximation to the distribution of undercount to result in an improvement in the overall distribution of population totals and shares.

The committee agrees that this assumption is one of the more vulnerable aspects of the PES design. Discussions and investigations of the homogeneity assumption occurred in the Undercount Steering Committee's deliberations prior to July 15, 1991, and the issue has been one of ongoing interest for the CAPE. Additional research has followed the report of the CAPE, inspired, in part, by comments received during the period of public comment.

Some reviewers provided only general remarks about their concerns with the homogeneity assumption, but others provided specific insights that spurred further investigation. One created a U.S. map showing the high degree of association between the adjustment at the state level and the groupings of states into the 4 census regions, North East, South, Midwest, and West, used in defining many of the post-strata. The reviewer showed maps of other characteristics, such as poverty rates, which do not exhibit so marked a regional character as the adjustments. Researchers at the Census Bureau subsequently reexamined the series of characteristics employed in defining the 8 sets of "artificial populations" - simulations of characteristics based on census data, such as the poverty rate, in a similar manner. To varying degrees, the Census Bureau's investigations confirm the point made by this reviewer, that is, that the adjustment methodology tends to emphasize regional aspects of the characteristic being estimated while missing or understating other components of state-to-state variation.

Another reviewer provided calculations showing that it was possible that departures from the homogeneity assumption, that is, heterogeneity, might account for more error in the PES adjustments of states than all the components of error estimated and included in the Census Bureau's total error model. This reviewer appeared to argue that a decision to adjust could not be reliably made when such a potentially large component of error had not been incorporated.

Consequently, research in this area has continued during the fall. The principal part of this research employed "artificial populations" based on actual population characteristics measured from the census, which were discussed in the previous report of the CAPE. Results included in the CAPE minutes of November 5, 1992 showed that, at the state level, the effect of departures from the homogeneity assumption tended generally to be large for the artificial populations. In fact, these investigations supported the strong possibility that the error due to heterogeneity could indeed be larger than all other sources of error in the adjustment, as one reviewer had suggested. In turn, the loss function analysis tended to understate the errors from adjustment, because heterogeneity bias tended to add to errors that the loss function analysis estimated for adjustment.

On the other hand, the investigations continued also to support the premise that the PES adjustment could still, on average, make improvements to the overall population shares. When heterogeneity bias is present, the results for artificial populations showed that the loss function analysis would tend to understate the errors of both the adjustment and of non-adjustment. If the loss function analysis understated the errors of adjustment and non-adjustment by equal amounts, then its estimate of the net difference would be correct. In fact, the analyses showed that, for a majority of the 8 populations, the loss function analysis would approximately correctly indicate or understate the net advantage from adjustment. There was an exception: the results for the artificial population based on unemployment gave unacceptable results, that is, the resulting loss function analysis would appear to have exaggerated substantially the net gains from adjustment.

2. Insufficient Sample Size

Some writers argued that the PES sample size was insufficient to permit an adjustment. These reviewers based their conclusion on comparisons to sizes of other samples with which they were familiar. These arguments were not reinforced, however, by explicit calculations showing in what sense the sample size was too small.

The issue of sample size is linked directly to the level of sampling variance, since increasing sample sizes predictably reduces sampling variance while not reducing most components of nonsampling error. The Census Bureau's total error model and loss function analysis were specifically designed to test whether the sample size was sufficiently large to obtain an improvement in the estimated total population and shares. In general, there is not one specific sample size that can be said to be large enough, since whether improvements can be made depends also on how much the undercount varies from one state to another, levels of measured nonsampling error, and the estimation procedure. The use of the 357-post-strata design reduced the effect of sampling variability considerably. By weighing the advantage that the PES would appear to accrue in correcting the census for large errors in states such as California against the small errors that would occur in estimating other states close to average undercount, the loss function analysis indicated that the PES sample size was sufficient to control uncertainty from sampling.

3. Post-Stratification

The Census Bureau decided to use a revised post-stratification scheme to control sampling variability instead of using a smoothing model. Several comments were received applauding this decision. However, some of these reviewers claimed that the post-stratification was data-driven. The end result of this was that the estimates of sampling error would be too low therefore causing the loss function analysis to unduly favor the adjustment.

One reviewer indicated that the revised post-stratification was acceptable, but indicated that a smoothing model would have been preferable. In addition, this reviewer indicated that an alternative technique to control sampling variability would have been to collapse the original post-stratification scheme based on

1392 categories. This would have had the effect of retaining greater homogeneity within post-strata. In the end, however, the reviewer (and the committee as well) felt that the revised 357 post-stratification scheme was superior to no adjustment.

These issues were discussed at various points by the committee. The committee was almost unanimous in deciding that smoothing would not be used in producing the revised post-stratified estimates. The committee was also pleased with the resulting post-stratification scheme. The committee recognized the danger of post-stratification, after data had been examined, and these concerns were documented in committee discussions.[1] This had some bearing on the general concerns that the committee expressed regarding the loss function analysis.

4. <u>Correlation Bias</u>

Correlation bias was widely discussed both internally and externally. Some reviewers generally noted that a correction based on correlation bias would be conservative in that it would not go far enough in correcting the differential undercount.

Other reviewers noted that at the national level there was clear evidence of correlation bias. However, they claimed that problems resulted because there were no direct measures of correlation bias sub-nationally. It was not clear to these reviewers that the methods of modeling correlation bias to produce sub-national estimates was appropriate, and there was concern that no supporting empirical evidence existed. Therefore, these reviewers were not convinced that the adjustment would improve the distribution of population shares sub-nationally.

The CAPE also expressed many of these same concerns as documented in the August 7, 1992 report, and in various meeting minutes.[2] The general conclusion of the committee was that correlation bias should be a component of total error. However, there were concerns expressed regarding the methods of estimating and allocating it. The committee requested that loss function analysis be done with and without correlation bias. Each committee member then had to make individual judgements about how to analyze the results.

---

[1]CAPE minutes from 9-18-91, 12-2-91, 12-30-91, 1-13-92, 1-27-92, 1-10-92, 1-16-92, 4-6-92, 4-22-92.

[2]CAPE minutes from 4-20-92, 6-1-92, 4-9-92, 6-11-92, 6-29-92, 9-1-92, 4-22-92

5. Total Error Model

Some reviewers viewed the total error model as being complete, and when combined with the loss function analysis supportive of an adjustment. One reviewer noted that he felt that the total error measurement of correlation bias was understated and a more accurate measurement would favor adjustment more than the current estimates.

There were other reviewers who did not believe that the total error model covered all sources of error. These reviewers cited various sources of error that they felt were omitted, such as: uncertainty from the choice of post-stratification or uncertainty from failure of the homogeneity assumption. These reviewers also felt that many of the sources of error included in the total error model were not measured accurately. They cited biases arising from imputation of missing data, fabrication error, and misreporting census day address as being particularly understated.

The CAPE discussed the total error model at great length.[3] The committee felt very confident that all components of error, except for bias due to failure of the homogeneity assumption, had been listed and considered. However, the committee could come to no agreement about the adequacy of the levels of error measured for the total error model components. The committee concluded in general to use caution in evaluating the results of the loss function analysis since the target numbers used in loss function analysis were so dependent on the levels of estimated bias.

6. Loss Function Analysis

Some reviewers viewed the loss function analysis as being very supportive of adjustment, and that the improvement indicated by the loss function analysis was an understatement (correlation bias was underestimated in the total error model).

Other reviewers generally had two major sources of concern regarding the loss function analysis: (1) There are uncertainties in the adjusted estimates that are not included in the loss function analysis, including uncertainty from failure of the homogeneity assumption, and from the choice of post-stratification. (2) There are concerns with the methods used to model the total error estimates of bias to create the target populations. In addition, one reviewer expressed concerns regarding the levels of significance reported for the loss function hypothesis tests.

---

[3]CAPE minutes from 4-20-92, 6-1-92, 6-9-92, 4-13-92, 4-22-92

The CAPE also discussed the loss function analysis in great detail.[4] In particular, the comments regarding uncertainty due to failure of the homogeneity assumption lead to some of the additional research reported in the addendum to the CAPE report. In general the committee accepted the loss function results keeping in mind a number of caveats.

7. Have All Questions Been Answered?

Many of the reviewers noted that there were many areas where additional research would provide useful information to inform the decision process. Some reviewers felt that sufficient information was available. Other reviewers generally, felt that more information was needed in order to justify an adjustment.

The August 7, 1992 report of the CAPE also indicates a number of areas where more research could be applied. Some of this research was continued and is reported in the addendum to the report. Many questions still remain. In spite of a desire for more complete information, the CAPE was able to reach general consensus that on average an adjustment to the 1990 census base would improve the intercensal estimates for states. For areas below the state level, the committee was able to reach no general consensus.

[4] CAPE minutes from 4-27-92, 5-4-92, 6-9-92, 6-24-92, 6-29-92, 7-6-92, 7-13-92, 9-1-92, 11-5-92, 4-22-92, 9-4-92

federal register

Monday
July 22, 1991

Part III

Department of Commerce

Office of the Secretary

Adjustment of the 1990 Census for Overcounts and Undercounts of Population and Housing; Notice of Final Decision

DEPARTMENT OF COMMERCE

Office of the Secretary

[Docket No. 91282-1181]

Decision of the Secretary of Commerce on Whether a Statistical Adjustment of the 1990 Census of Population and Housing Should Be Made for Coverage Deficiencies Resulting in an Overcount or Undercount of the Population

AGENCY: U.S. Department of Commerce.

ACTION: Notice of final decision.

SUMMARY: This is a notice of the final decision of the Secretary of Commerce on the issue of adjusting the 1990 census to correct for overcounts or undercounts of the population in the 1990 Decennial Census. The purpose of this notice is to inform the public of the decision and to explain the basis for the decision.

DATES: The decision is effective on July 15, 1991.

FOR FURTHER INFORMATION CONTACT: Michael R. Darby, Under Secretary for Economic Affairs and Administrator, Economics and Statistics Administration, Room 4848 Herbert C. Hoover Building, United States Department of Commerce, Washington, DC 20230, Telephone (202) 377-3727.

SUPPLEMENTARY INFORMATION: The Secretary of Commerce is required, pursuant to 13 U.S.C. 141, to conduct a decennial census of the population of the United States. The population totals derived from the census provide the basis for the apportionment of seats in the United States House of Representatives, for state legislative redistricting, for determining district boundaries for county and city elections, and for the allocation of federal funds to state and local governments.

In 1987, the Secretary of Commerce decided not to plan for a statistical adjustment of the 1990 census. As a result, a lawsuit was filed by the city of New York and other parties seeking to compel the Department to plan for such an adjustment. Pursuant to an agreement between the parties in *City of New York, et al.* v. *Department of Commerce, et al.*, 88-Civ.-3474 (E.D.N.Y.), the Department undertook a *de novo* review of the adjustment issue in order to make a decision no later than July 15, 1991, on whether to adjust the 1990 census. The purpose of this notice is to inform the public about the Secretary's decision and the basis for the decision.

Final guidelines which aided the Secretary in his decision were published in the Federal Register on March 15, 1990 (FR vol. 55, no. 51, part III pp. 9838-9861).[1] They were intended to provide the framework for a balanced consideration by the Secretary of factors relevant to the decision.

The census adjustment decision process was divided into several distinct phases. The first phase was the actual enumeration of the population. The second phase was the conduct of a post-enumeration survey, based on a probability sample of housing units. This sample provided data for two purposes: estimation of the net overcount or undercount of basic enumeration subgroups using capture-recapture methodology, and application of factors for the adjustment of the enumerated counts. The third phase of the process was a determination of the adequacy of the post-enumeration survey as an evaluation and adjustment tool. The fourth and final phase of the process was a decision on the adjustment question by the Secretary based on the published guidelines.

In making his decision, the Secretary relied on the advice of senior officials in the Economics and Statistics Administration, which includes the Census Bureau, as well as other senior advisors. The Secretary also relied on the individual recommendations of the eight members of the Special Advisory Panel appointed to provide independent advice to the Secretary on the adjustment question. In addition, the Secretary considered the public comments submitted to the Department pursuant to a Federal Register notice dated May 24, 1991, seeking comments on the question of whether the 1990 Census should be adjusted. The Department received approximately 650 public comments. These comments, as well as the appendices referred to in the following explanation of the decision, are available for public inspection in the U.S. Department of Commerce Central Reference and Records Inspection Facility, room 6020 Herbert C. Hoover Building, 14th Street and Constitution Avenue, NW., Washington, DC 20230.

Following is a detailed discussion of the adjustment decision and the basis for the decision. The discussion is in four sections: a summary statement, an analysis of the guidelines, an evaluation of the recommendations of the Special Advisory Panel and a statement of the decennial census procedures.

[1] Proposed guidelines were published in the Federal Register on December 11, 1989. The Court has previously considered and rejected a challenge to the guidelines. See *City of New York v. United States Department of Commerce*, 739 F.Supp. 767 (E.D.N.Y. 1990).

Dated: July 15, 1991.

Robert A. Mosbacher,

Secretary of Commerce.

SECTION 1—SUMMARY STATEMENT

Statement of Secretary Robert A. Mosbacher on Adjustment of the 1990 Census

Reaching a decision on the adjustment of the 1990 census has been among the most difficult decisions I have ever made. There are strong equity arguments both for and against adjustment. But most importantly, the census counts are the basis for the political representation of every American, in every state, county, city, and block across the country.

If we change the counts by a computerized, statistical process, we abandon a two hundred year tradition of how we actually count people. Before we take a step of that magnitude, we must be certain that it would make the census better and the distribution of the population more accurate. After a thorough review, I find the evidence in support of an adjustment to be inconclusive and unconvincing. Therefore, I have decided that the 1990 census counts should not be changed by a statistical adjustment.

The 1990 census is one of the two best censuses ever taken in this country. We located about 98 percent of all the people living in the United States as well as U.S. military personnel living overseas, which is an extraordinary feat given the size, diversity and mobility of our population. But I am sad to report that despite the most aggressive outreach program in our nation's history, census participation and coverage was lower than average among certain segments of our population. Based on our estimates, Blacks appear to have been undercounted in the 1990 census by 4.8%, Hispanics by 5.2%, Asian-Pacific Islanders by 3.1%, and American Indians by 5.0%, while non-Blacks appear to have been undercounted by 1.7%.

I am deeply troubled by this problem of differential participation and undercount of minorities, and I regret that an adjustment does not address this phenomenon without adversely affecting the integrity of the census. Ultimately, I had to make the decision which was fairest for all Americans.

The 1990 census is not the vehicle to address the equity concerns raised by the undercount. Nonetheless, I am today requesting that the Census Bureau incorporate, as appropriate, information gleaned from the Post-Enumeration Survey into its intercensal estimates of the population. We should also seek other avenues for the Bush Administration and Congress to work together and address the impact of the differential undercount of minorities on federal programs.

In reaching the decision not to adjust the census, I have benefitted from frank and open discussions of the full range of issues with my staff, with senior professionals from the Economics and Statistics Administration and the Census Bureau, with my Inspector General, and with statisticians and other experts. Throughout these discussions, there was a wide range of professional opinion and honest disagreement. The Department has tried to make the process leading to this decision as open as possible. In that spirit, we will provide the full record of the basis for our decision as soon as it is available.

In reaching the decision, I looked to statistical science for the evidence on whether the adjusted estimates were more accurate than the census count. As I am not a statistician, I relied on the advice of the Director of the Census Bureau, the Associate Director for the Decennial Census and other career Bureau officials, and the Under Secretary for Economic Affairs and Administrator of the Economics and Statistics Administration. I was also fortunate to have the independent counsel of the eight members of my Special Advisory Panel. These eight experts and their dedicated staffs gave generously of their time and expertise, and I am grateful to them.

There was a diversity of opinion among my advisors. The Special Advisory Panel split evenly as to whether there was convincing evidence that the adjusted counts were more accurate. There was also disagreement among the professionals in the Commerce Department, which includes the Economics and Statistics Administration and the Census Bureau. This compounded the difficulty of the decision for me. Ultimately, I was compelled to conclude that we cannot proceed on unstable ground in such an important matter of public policy.

The experts have raised some fundamental questions about an adjustment. The Post-Enumeration Survey, which was designed to allow us to find people we had missed, also missed important segments of the population. The models used to infer populations across the nation depended heavily on assumptions, and the results changed in important ways when the assumptions changed. These problems don't disqualify the adjustment automatically—they mean we won't get a perfect count from an adjustment. The question is whether we will get better estimates of the population. But what does better mean?

First, we have to look at various levels of geography—whether the counts are better at national, state, local, and block levels. Secondly, we have to determine both whether the actual count is better and whether the share of states and cities within the total population is better. The paradox is that in attempting to make the actual count more accurate by an adjustment, we might be making the shares less accurate. The shares are very important because they determine how many congressional seats each state gets, how political representation is allocated within states, and how large a "slice of the pie" of federal funds goes to each city and state. Any upward adjustment of one share necessarily means a downward adjustment of another. Because there is a loser for every winner, we need solid ground to stand on in making any changes. I do not find solid enough ground to proceed with an adjustment.

To make comparisons between the accuracy of the census and the adjusted numbers, various types of statistical tests are used. There is general agreement that at the national level, the adjusted counts are better, though independent analysis shows that adjusted counts, too, suffer from serious flaws. Below the national level, however, the experts disagree with respect to the accuracy of the shares measured from an adjustment. The classical statistical tests of whether accuracy is improved by an adjustment at state and local levels show mixed results and depend critically on assessments of the amount of statistical variation in the survey. Some question the validity of these tests, and many believe more work is necessary before we are sure of the conclusions.

Based on the measurements so far completed, the Census Bureau estimated that the proportional share of about 29 states would be made more accurate and about 21 states would be made less accurate by adjustment. Looking at cities, the census appears more accurate in 11 of the 23 metropolitan areas with 500,000 or more persons: Phoenix, Washington, DC, Jacksonville, Chicago, Baltimore, New York City, Memphis, Dallas, El Paso, Houston and San Antonio. Many large cities would appear to be less accurately treated under an adjustment. While these analyses indicate that more people live in jurisdictions where the adjusted counts appear more accurate, one third of the population lives in areas where the census appears more accurate. As the population units get smaller, including small and medium sized cities, the adjusted figures become increasingly unreliable. When the Census Bureau made allowances for plausible estimates of factors not yet measured, these comparisons shifted toward favoring the accuracy of the census enumeration. Using this test, 28 or 29 states were estimated to be made less accurate if the adjustment were to be used. What all these tests show, and no one disputes, is that the adjusted figures for some localities will be an improvement and for others the census counts will be better. While we know that some will fare better and some will fare worse under an adjustment, we don't really know how much better or how much worse. If the scientists cannot agree on these issues, how can we expect the losing cities and states as well as the American public to accept this change?

The evidence also raises questions about the stability of adjustment procedures. To calculate a nationwide adjustment from the survey, a series of statistical models are used which depend on simplifying assumptions. Changes in these assumptions result in different population estimates. Consider the results of two possible adjustment methods that were released by the Census Bureau on June 13, 1991. The technical differences are small, but the differences in results are significant. The apportionment of the House of Representatives under the selected scheme moved two seats relative to the apportionment implied by the census, whereas the modified method moved only one seat. One expert found that among five reasonable alternative methods of calculating adjustments, none of the resulting apportionments of the House were the same, and eleven different states either lost or gained a seat in at least one of the five methods. I recognize that the formulas for apportioning the House are responsive to small changes and some sensitivity should be expected. What is unsettling, however, is that the choice of the adjustment method selected by Bureau officials can make a difference in apportionment, and the political outcome of that choice can be known in advance. I am confident that political considerations played no role in the Census Bureau's choice of an adjustment model for the 1990 census. I am deeply concerned, however, that adjustment would open the door to political tampering with the census in the future. The outcome of the enumeration process cannot be directly affected in such a way.

My concerns about adjustment are compounded by the problems an adjustment might cause in the redistricting process, which is contentious and litigious enough without an adjustment. An adjusted set of numbers will certainly disrupt the political process and may create paralysis in the states that are working on redistricting or have completed it. Some people claim that they will be denied their rightful political representation without an adjustment. Those claims assume that the distribution of the population is improved by an adjustment. This conclusion is not warranted based on the evidence available.

I also have serious concerns about the effect an adjustment might have on future censuses. I am worried that an adjustment would remove the incentive of states and localities to join in the effort to get a full and complete count. The Census Bureau relies heavily on the active support of state and local leaders to encourage census participation in their communities. Because census counts are the basis for political representation and federal funding allocations, communities have a vital interest in achieving the highest possible participation rates. If civic leaders and local officials believe that an adjustment will rectify the failures in the census, they will be hard pressed to justify putting census outreach programs above the many other needs clamoring for their limited resources. Without the partnership of states and cities in creating public awareness and a sense of involvement in the census, the result is likely to be a further decline in participation.

In looking at the record of public comment on this issue, I am struck by the fact that many civic leaders are under the mistaken impression that an adjustment will fix a particular problem they have identified—for example, specific housing units or group quarters that they believe we missed. It does not do so. It is not a recount. What an adjustment would do is add over 6 million unidentified people to the census by duplicating the records of people already counted in the census while subtracting over 900,000 people who were actually identified and counted. The decisions about which places gain people and which lose people are based on statistical conclusions drawn from the sample survey. The additions and deletions in any particular community are often based largely on data gathered from communities in other states.

The procedures that would be used to adjust the census are at the forefront of statistical methodology. Such research deserves and requires careful professional scrutiny before it is used to affect the allocation of political representation. Since the results of the evaluation studies of the survey were made available, several mistakes have been found which altered the certainty of some of the conclusions drawn by my advisors. The analysis continues, and new findings are likely. I am concerned that if an adjustment were made, it would be made on the basis of research conclusions that may well be reversed in the next several months.

It is important that research on this problem continue. We will also continue the open discussion of the quality of the census and the survey and will release additional data so that independent experts can analyze it. We must also look forward to the next census. Planning for the year 2000 has begun. A public advisory committee on the next census has been established and by early fall I will announce the membership of that committee. I have instructed the Census Bureau's Year 2000 task force to consider all options for the next census, including methods for achieving sound adjustment techniques.

I give my heartfelt thanks to the many people who have devoted so much time and energy to this enterprise. The staff at the Census Bureau have demonstrated their professionalism at every turn through the last two difficult years. They executed a fine census and an excellent survey and then condensed a challenging research program into a few short months. I am deeply grateful for their help. Let me reiterate my sincere thanks to the Special Advisory Panel for their substantial contribution. The staff at the Department, especially those in the Economics and Statistics Administration, also deserve praise.

With this difficult decision behind us, we will commit ourselves anew to finding sound, fair and acceptable ways to continue to improve the census process. We welcome the leadership of Congress and other public officials, community groups, and technical experts in maximizing the effectiveness and minimizing the difficulties of the year 2000 census.

July 15, 1991.

SECTION 2—ANALYSIS OF THE GUIDELINES

Analysis of the Guidelines

Introduction

The 1990 census counts should not be changed by a statistical adjustment. This section explains my evaluation of the evidence relevant to each of the eight guidelines that I considered in reaching my decision. Each section begins with a statement of the guideline and a reiteration of the explanation of the guideline contained in the March 15, 1990, Federal Register notice. A discussion of the guideline follows. The final section states my conclusions.

Summaries of my conclusions on each of the eight guidelines are set forth below.

Guideline One

Guideline One requires that convincing evidence be offered that the adjusted estimates of the population are more accurate than the census at the national, State, and local levels. In the absence of such evidence, the census counts are concluded to be the most accurate.

At the national level, it is likely that the PES-adjusted estimates reflect more accurately the total population and the racial and ethnic populations of the country. It appears equally clear, however, that the PES omitted large numbers of certain groups—notably black males. We have no information on the location of these persons. In addition, the PES and demographic analysis lead to sharply different conclusions about the accuracy of the census for several age/sex groups at the national level. Although these are not definitive disqualifiers at the national level, they do raise some question as to whether the adjusted figures are more accurate than the census count even at the national level.

The Constitution requires a census every 10 years not just to count the total number of people in the United States but to locate them so that political representation can be allocated to the states and the people in them in proportion to their numbers. I conclude that the primary criterion for accuracy should be distributive accuracy—that is, getting most nearly correct the proportions of people in different areas. Improved numeric accuracy, although in itself desirable, cannot compensate for treating states and individuals less fairly.

At the State and local level the correct statistical analysis for both distributive and numeric accuracy simply has not been completed. The total error model indicates that the adjusted figures tend to be too high but generally closer in numeric terms to the true population than the census counts which tend to be too low. However, there is sufficient uncertainty about the true variance of the adjusted figures that even numeric accuracy has not been definitively demonstrated. The loss function analysis and hypothesis tests that have been prepared by the Census Bureau to date, although of uncertain reliability, do support the superior accuracy of the census counts versus the adjusted figures when we consider distributive accuracy—or fairness—and use reasonable estimates of the error variance of the alternative DSE. That is, for the Constitutional purposes of the census the available evidence is consistent with the census counts being more accurate than the adjusted counts. There is certainly not sufficient evidence to reject the distributive accuracy of the census counts in favor of the adjusted counts.

I conclude that, in accordance with Guideline One, the census counts are the most accurate count of the population of the United States at the State and local levels. While the preponderance of the evidence leads me to believe that the total population at the national level falls between the census counts and the adjusted figures, that conclusion is not relevant to the determination of distributive accuracy. Thus this guideline weighs in favor of a decision not to adjust.

Guideline Two

I conclude that the considerations pointed to by Guideline Two tend to reject use of the adjusted figures and support use of the census counts. The adjusted figures—like the census counts—are consistent across all jurisdictional levels and of sufficient detail for all purposes. However, the adjusted figures do not appear to be of sufficient quality to be usable for reapportionment and redistricting. First, the distributive accuracy of the census counts is superior as concluded above in my review of the evidence on Guideline One. Furthermore, substantial evidence casts doubt on the homogeneity assumption underlying the entire synthetic adjustment methodology. Even if the tests discussed under Guideline One and based on the homogeneity assumption had proven favorable to adjustment, this evidence would weigh against adjustment. Instead, both considerations imply that the adjusted figures are not of sufficient quality to be usable for reapportionment and legislative redistricting. Thus, this Guideline weighs in favor of a decision not to adjust the census.

Guideline Three

I have previously concluded that the adjusted figures have not been shown to be more accurate than the census enumeration. That is all that is required under Guideline Three to conclude that the census may not be adjusted. There are, however, additional considerations under Guideline Three under which I also conclude the 1990 census should not be adjusted.

It has proved virtually an impossible task to prespecify the adjustment procedure. It is equally impossible to prespecify the Census procedure. However, in the adjustment procedure an individual or responsible group must make choices which have politically significant effects on the counts that can be transparent to those making the choices. This puts the counts at greater risk of being manipulated than the census. There is no evidence of unprofessional or political manipulation in the 1990 PES program.

The results of the adjustment procedure are broadly robust at an aggregate, national level. However, although various alternatives seem to distribute counts in roughly similar ways, small changes in methodology can move seats in the House. It is also true that small changes in the census enumeration can move seats in the House as well, but no individual involved in the enumeration process can predict how. That is not true for the decisions for adjustment that cannot be or were not prespecified.

One of the most problematic parts of the adjusted process was the bundle of statistical techniques contained in the smoothing process. These techniques relied heavily on statistical assumptions, resulted in large changes in adjustment factors, and may very well have led to an overstatement of the undercount. Thus, this guideline weighs in favor of a decision not to adjust.

Guideline Four

Based on the information available, I conclude that an adjustment would adversely affect future census efforts to a greater extent than any adverse effects of a decision not to adjust. The evidence indicates that the controversy over adjustment is likely to have a negative effect on future censuses regardless of the outcome of the adjustment decision. I am concerned that an adjustment would reduce state and local support for future censuses, adversely affect the Department's ability to obtain appropriate funding for future censuses, adversely affect the quality of the work done in the future by temporary census enumerators who are essential in reaching the hard-to-count, subject the Census Bureau to partisan pressures, and create the possibility for political manipulation of future census counts. Thus, this guideline weighs in favor of a decision not to adjust.

Guideline Five

The question whether the chosen method of adjustment would violate the Constitution and federal statutes depends upon the substantive analysis of whether accuracy of the census is improved by an adjustment. Because there are other compelling substantive reasons not to adjust, legal considerations did not provide a basis for my decision.

Guideline Six

An adjustment to the census is a fundamental change in the way we count and locate the persons residing in the United States. I am deeply concerned that if an adjustment is made, it would be made on the basis of research conclusions that may very well be reversed in the next several months. That would be bad for the country and bad for the Census Bureau.

The results of the PES evaluation studies are not yet completely analyzed. Because of the compressed time schedule imposed by the July 15 deadline, the analysis has not been subject to the full professional scrutiny that such important research requires and deserves. To the Census Bureau's great credit, the statistical tools used to calculate and evaluate the adjusted counts are at the cutting edge of statistical research. But such cutting edge research is not tried and true—it requires more thorough scrutiny before it can be used to affect the allocation of political representation and Federal funding.

Nonetheless, the demands of good research must be weighed against the need for a timely decision. In time we may find a way of combining the PES and the census to create counts that better reflect the absolute levels and the distribution of the population. There are sufficient data and analysis to support a decision not to adjust.

Guideline Seven

Any decision will result in some level of disruption through legal challenges. On balance, the record indicates that a decision to adjust would likely be more disruptive than a decision not to adjust. A decision to adjust would clearly cause disruption in those States that have early redistricting deadlines. The assertion that persons are denied their rightful claims without an adjustment assumes that the distribution of the population is improved by an adjustment. Based on the evidence, this assumption is invalid. Thus, this guideline weighs in favor of a decision not to adjust.

Guideline Eight

The requirements for this Guideline been met. This Guideline does not h in favor of a decision either way since the requirements of this Guideline could have been fully met if the decision had been to adjust.

Guideline One

The Census shall be considered the most accurate count of the population of the United States, at the national, State and local level, unless an adjusted count is shown to be more accurate. The criteria for accuracy shall follow accepted statistical practice and shall require the highest level of professional judgment from the Bureau of the Census. No statistical or inferential procedure may be used as a substitute for the Census. Such procedures may only be used as supplements to the Census.

Explanation

The mandate of the Census Bureau is to enumerate the population in a manner that assures that the count of the population is the best achievable given current methodology. As stated in the introduction, the assertion that a method involving statistical inference could lead to a more accurate enumeration warrants close scrutiny.

et of adjusted counts would be on a statistical inference that counted for persons were present and that persons who were actually enumerated do not exist or were counted twice. Both determinations are based on a survey of a sample of similar blocks from locations across the country. Thus, the evidence, to be acceptable, must show convincingly that the count can be improved by statistical adjustment at national, state and local levels. In making this assessment, we will examine the effects of the proposed adjustment on the accuracy of counts at all geographic levels.

Comparison of estimates of population size. The estimates of the size of the population from the original enumeration, the demographic analysis, and the post-enumeration-survey estimates will be compared to assess their consistency. The comparison will take into consideration the uncertainty inherent in the demographic analysis and post-enumeration-survey estimates. For the reasons explained in the introduction, the original enumerations will be considered to be more accurate for all geographic areas unless the evidence from demographic analysis and the post-enumeration survey strates convincingly that the stem estimate is more accurate.

Accordingly, the Bureau of the Census shall carefully scrutinize and fully describe the size of any net undercount or net overcount inferred from demographic analyses of population sub-groups and the sources of any net undercount or net overcount of population subgroups inferred from the analysis of the post-enumeration survey.

Technical Grounds

Demographic Analysis. Estimates of the size of certain cohorts of the population are based on assumptions about or studies of the behavior of those populations. For some cohorts these assumptions have led to conclusions of net undercounts or net overcounts in several different censuses. The extent to which such conclusions result from specific assumptions will be described. Moreover, the extent to which these assumptions are warranted, and the sensitivity of such conclusions to changes in these assumptions, will be assessed.

The potential sources of error in the demographic analyses the Bureau currently plans are:

Birth registration completeness.
Net immigration of undocumented aliens.
White births, 1915–1935.
Black births, 1915–1935.
Foreign-born emigrants.
Population over age 65.
Models to translate historical birth-record racial classifications into 1990 self-reported census concepts.

The Bureau will examine the effect of errors in each of these measurements on estimates of the net overcount or net undercount. These studies will yield ranges of uncertainty for the demographic estimates of the population which will in turn yield ranges of uncertainty for the net overcount or net undercount. The effect of uncertainty in each of these components will be cumulated into overall levels of potential error.

Post-Enumeration Survey. The capture-recapture method lies at the heart of the post-enumeration-survey models for estimating population coverage deficiencies. The use of this methodology to derive the net undercount or net overcount estimates will be clearly explained. The appropriateness of this methodology to the enumeration of the population will be assessed.

Like demographic analysis, the post-enumeration-survey adjustment mechanism relies on numerous assumptions. The extent to which these assumptions are warranted, and the sensitivity of the conclusions to changes in these assumptions, will be assessed.

Survey methods are based on randomly chosen samples that use statistical inference to estimate the population of the Nation and its components. Such estimates are subject to statistical variation within some range of values—that is, a replication of the process used to make the estimate (including taking the sample) may not lead to the same estimate as the original procedures. Thus, there is a likely range of estimates around the "true" count of the population that depends on the random sample chosen.

If the range of estimates likely to occur is small and near the "truth," then any particular estimate is close to the truth and, thus, acceptable as an approximation of the "truth." If the range is very large, then any particular estimate may not be close to the "truth," and the estimation process gives us little information about the "truth."

A relevant technical criterion related to uncertainty introduced by sampling is how small any possible range of dual-system estimates must be to conclude that any particular outcome of the dual-system estimation process is more accurate than the enumeration itself.

Because the post-enumeration survey itself is a sample, the quantified parameters of the deficiencies are themselves estimates and subject to statistical variability. This variability must be small enough to ensure that any modification of the enumeration is an improvement over the unadjusted counts.

The post-enumeration survey serves two functions. The first function is to detect any deficiencies in the enumeration. For the post-enumeration survey to show convincingly that the enumeration is deficient, it must be clear that the deficiencies are not a result of problems in taking the post-enumeration survey. It follows, then, that the quality of the post-enumeration survey is a central concern in the decision whether to adjust.

The second function is to quantify any deficiencies attributed to the enumeration precisely enough to allow the enumeration to be modified in such a way that we are reasonably certain that the modified enumeration is more accurate than the original enumeration. Thus the post-enumeration survey must quantify the deficiencies of the enumeration precisely and accurately.

How much uncertainty in the measures of deficiency of the enumeration is acceptable?

(1) If the likely range of measures of deficiency would include outcomes that would call for no modification in the

enumeration, then no modification would be done.

(2) The enumeration could be modified if the likely range of measures of deficiency would lead to potential modifications that would be substantially similar in terms of their impact on the counts of demographic groups, their impact on apportionment of Congress, and their impact on local population counts.

The quality of the net overcount or net undercount estimates that result from the post-enumeration survey depends on the quality of a series of operations used to gather and process the required data. The Bureau of the Census will undertake a series of studies to assess the statistical quality of the post-enumeration survey data. The results of these studies will yield measures of the precision and accuracy of the net overcount and net undercount estimates and a range of estimates for the net undercount and net overcount.

The current plans of the Bureau include investigation of the following sources of error for the dual system estimate of population size based on the post-enumeration survey and the census:

Missing data
Quality of the reported census day address
Fabrication in the P sample
Matching error
Measurement of erroneous enumerations
Balancing the estimates of gross overcount and gross undercount
Correlation bias
Random error

These and other component errors will be combined to produce an estimate of the overall level of error. In all evaluations, analyses will examine data for the population as a whole and for race, sex, Hispanic origin, and geographical detail.

Discussion

To certify a set of adjusted counts as the official counts of the population of the United States, one must accept the statistical inferences from a survey that there are persons who were unaccounted for by the census but who were actually present in a specific location on census day, that persons who were actually enumerated either did not exist or were counted twice, and that the same survey, when combined with census counts, can produce more accurate figures than the census enumeration alone. All these inferences are based on information from a sample of 377,381 persons in 171,390 housing units and group quarters in 5,290 block clusters. The people who are inferred to be missing from the census or erroneously enumerated in the census must then be correctly allocated to the specific blocks in which these mistakes were made. These blocks must be chosen out of the 4,830,514 inhabited blocks in the United States. Thus, acceptance of adjusted counts as more accurate requires not only that the counts themselves be shown to be more accurate, but that the distribution of those counts across the United States reflect more accurately the distribution of the population. This is the burden of proof imposed by Guideline One on any decision to adjust the census.

There are three population measurement techniques that play a role in making these statistical inferences. The first is the census enumeration. This was an effort to count each and every person residing in the United States on April 1, 1990. The second is the Post-Enumeration Survey (PES). This is the survey mentioned in the preceding paragraph that was taken several months after census day, independently of the census. An attempt is made to match the persons surveyed in the PES back to records in the census and to match persons in the census to the PES. From the results of this matching process, and a complex web of statistical models, inferences can be made about the number of persons missed by the census and their location. It is the quality of these inferences that is at issue. The third technique is called demographic analysis (DA). DA makes an independent estimate of the population at a national level from administrative records. It can be used to calibrate the results from the census or PES. DA calculates the population from the number of births, number of deaths, the number of immigrants, and the number of emigrants. It builds up a count of the population of the United States from birth and death certificates, immigration records and other sources. Like the census and the PES, DA is also an imperfect measure, so the quality of the inferences made from it are in question as well.

In the course of the discussion of this guideline, various aspects of these three complex processes will be explained and discussed. A detailed explanation can be found in Section Four of this report. We begin by comparing the national counts found in 1990 using these three methods.

A Comparison of the Counts at the National Level Using Three Methods

The national total count from the census enumeration is compared, in Table 1, Appendix 14, with the corresponding total in the proposed adjusted counts based on the PES and also with the corresponding estimates based on DA. The census count is 2.07% or 5,269,917 persons less than the PES estimate. There is evidence of racial, ethnic, and sex differential undercounts in the census when compared to the PES-based estimates. The count of black males in the census was 5.37% or 804,233 persons lower than the population inferred from the PES. The count of black females in the census was 4.33% or 715,543 persons lower than the PES estimated. For non-black males the census count was below the PES estimate by 2.02% or 2,205,443 persons and for non-black females the differential was 1.36% or 1,544,050 persons.

Estimates of national population totals are derived by DA based primarily on administrative records. Demographic analysis estimates provide national totals only and cannot be used to locate people as census counts are required to do. Many argue that the DA estimates broadly corroborate differential undercounts implied by PES-adjusted counts; [1] however, like the minority on the Undercount Steering Committee,[2] I find there are some important and puzzling differences. First, the overall undercount rate inferred from comparing the census to DA (1.85%) is smaller than that inferred from the PES (2.07%). At an aggregate level, the demographic analysis is thought to be more inclusive since the PES and census will miss people who are difficult to survey. Thus the estimate of the population from the PES was expected to be lower than the DA estimate. It is not. The PES estimated total population is 0.23% higher than the DA estimate. More detailed analysis shows that the PES and DA estimates are not far apart in a statistical sense.[3]

[1] See appendix 7: Bryant, Barbara E., Director of the U.S. Bureau of the Census, "Recommendation to Secretary of Commerce Robert A. Mosbacher on Whether or Not to Adjust the 1990 Census," June 28, 1991. [hereafter Bryant] page 18. See also Appendix 4, "Report of the Undercount Steering Committee," U.S. Bureau of the Census, June 21, 1991. [hereafter Undercount Steering Committee] page 4. See also Appendix 8: Ericksen, Eugene P., Estrada, Leobardo F., Tukey, John W., Wolter, Kirk M. "Report on the 1990 Decennial Census and the Post-Enumeration Survey." Members of the Special Advisory Panel, June 21, 1991. [hereafter Ericksen, *et al.*] page 10.

[2] Undercount Steering Committee, page 4.

[3] The 95% confidence interval for the overall PES undercount rate is from 1.83% to 2.30% and the judgmental 95% confidence interval for the overall demographic undercount rate is from 1.6% to 2.4%. A confidence interval gives the range of statistically plausible values. The "95%" refers to the notion that one is 95% sure this interval has captured the true, but unknown, value. See table 2 in appendix 14.

Nevertheless, the fact that the direction of difference is the opposite of what ...istical experience would have led us ...pect raises a troubling question ...t the relationship between the two methods.[4]

Another example of a gross inconsistency between the PES and DA is that an adjustment would add 1,055,828 more females than DA indicates should be added. If DA were in fact correct, and the enumeration were adjusted, the official population counts would have a 0.82% overcount of females imbedded in it.

The third disturbing comparison between the PES and DA undercount rates is that all groups of black males (except those aged 10–19) are substantially undercovered by the PES relative to DA. This results in PES-based undercount rates that are substantially smaller than the DA rates. This is the type of result that is usually expected in comparing the PES and DA.[5] An adjustment based on the PES would add 804,233 black males to the population. According to demographic analysis, the number of black males that should be added to the population is 1,338,380. Thus the PES-based adjustment would be omitting 534,147 black males according to DA. For black females the PES adjustment would add 29,390 fewer persons than DA indicates should be ...d. If we accept the DA as being ...r to the truth, we could not ...opriately add the persons the PES missed to the count because we have no way of locating them.

Some will argue that "going part way" toward remedying the undercount of black males is better than doing nothing.[6] The trouble with this argument is that it ignores the fact that increased accuracy for census counts means not only increased accuracy in the level of the population, but also increased accuracy in the distribution of the population in states and localities. In particular, for the primary uses of the census—apportionment and redistricting—the share or fraction of the total population in a given state, city or precinct is critical. It is this fraction that determines political representation and the amount of Federal funds allocated across political jurisdictions. The paradox is that even if you improve the accuracy in the level of the population in any given city by adding at least some of the people missed in the census, you do not necessarily improve and can worsen accuracy in the share of the population in that city. This point is explored further in the section on how accuracy is measured.

Special Advisory Panel Member Wachter estimates that the number of people missed by both the census and the PES may be as high as half-a-million.[7] We do not know where these people are.[8] The implicit assumption that we would be making if we went ahead and adjusted the count is that they are spread over the country in the same way as the post-adjustment population. Such an assumption has no empirical foundation. There is no doubt that there is a fundamental deficiency in the count, but there is also a fundamental deficiency in the PES. It is not clear that the adjusted counts will accurately reflect relative populations in particular jurisdictions. As Wachter states:

> When we try to gauge the relative sizes of two states or cities or counties or districts [after an adjustment], we must always worry that there are enough more of the unreached in one than in the other to reverse the judgment about relative size that the adjusted counts would lead us to make.[9]

To further complicate matters, at the national level there are instances where a PES-based adjustment to the census would move subpopulation totals in the opposite direction from that indicated by DA:

- An adjustment based on the PES will add 180,318 non-black males aged 10–19, while the DA indicates 136,908 should be deleted, a difference in the wrong direction of 317,226.[10]
- An adjustment based on the PES will delete 91,631 males over the age of 65, while DA indicates that 192,930 should be added, a difference in the wrong direction of 284,541 persons.[11]
- An adjustment based on the PES will add 375,053 females aged 10–19 when DA indicates that 7,141 should be deleted, a difference of in the wrong direction of 382,191.[12]
- An adjustment based on the PES will delete 245,253 females over the age of 45 while DA indicates 146,255 should be added, a difference of 391,508 persons in the wrong direction.[13]

Another grouping of the population that plays a key role in the adjustment process is called a post-stratum. To calculate the adjusted population estimates, the population is broken down into 1392 groups called post-strata. Every individual in the United States fits into one, and only one, of these post-strata. These post-strata are based on census division, type of place of residence, tenure of residence, race, Hispanic ethnicity, sex, and age. These are the smallest groupings of people for which an undercount rate is estimated by the Census Bureau. When post-strata for similar types of persons are combined (for example, all post-strata with blacks, or all post-strata for people age 30–44) the results are largely consistent with expectations.[14] However, there is a lot of variation across the post-strata for similar types of people. Wachter offers intriguing evidence that "the story of census coverage, at a level of fine detail, is more complicated than one would hope."[15] For example, if one looks at all the post-strata for blacks, 25% of them show an overcount rather than an undercount.[16] Thus the broad, national-level aggregations of undercount by race, ethnicity, sex, and age mask a large amount of diversity within those groups. It is therefore overly simplistic to conclude that the census generally results in an undercount for all members of any particular group.

---

[4] As will be discussed later, there are measured biases in the production adjustment estimates. When corrections are made for these measured biases, the overall undercount rate measured by the PES falls below that of DA.

[5] The technical term for this is correlation bias.

[6] See Undercount Steering Committee, page 4; See also appendix 8: Ericksen, Eugene P. "Recommendation on 1990 Census Adjustment," Member, Special Advisory Panel, June 21, 1991, [hereafter Ericksen] page 2; See also Appendix 8: Estrada, Leobardo F. "Recommendation on 1990 Census Adjustment," Member, Special Advisory Panel, June 21, 1991, [hereafter Estrada] page 14; See also Appendix 8: Wolter, Kirk W. "Recommendation on 1990 Census Adjustment," ...ber, Special Advisory Panel, June 21, 1991, ...fter Wolter] page 4.

[7] See appendix 8: Wachter, Kenneth W. "Recommendations on 1990 Census Adjustment," Member, Special Advisory Panel, June 17, 1991, [hereafter Wachter] page 8.

[8] The implications of this for accuracy are explained at length below.

[9] Wachter, page 8.

[10] The third table in appendix 14 shows that the 95% PES confidence interval for the undercount rate for this group is (0.53, 1.85) with a point estimate of 1.19. Demographic analysis shows a confidence range of (−1.21, 0.65) with a point estimate of −0.82. Thus neither estimate falls in the other's confidence range.

[11] The third and fourth table of appendix 14 show the confidence intervals for undercount rates for blacks and non-blacks separately. For non-blacks in this group, the confidence intervals for the two methods do not intersect, with the PES confidence interval completely less than zero and the DA confidence interval completely greater than zero. For blacks as well, the two intervals do not overlap. The PES spans zero, the DA is completely greater than zero.

[12] In appendix 14 the confidence intervals for this group are given for blacks and nonblacks separately. For non-blacks the intervals for the PES and DA do not overlap. For blacks they do.

[13] The confidence intervals for the four component groups are given in tables 1 and 2 of appendix 14. The intervals are wide enough that the differences may not be statistically significant.

[14] Wachter, pages 9–10.

[15] Wachter, page 10.

[16] Wachter, page 10.

This section has given an aggregate picture of the population using three different measurement instruments—the census, the PES, and DA. It is clear that the census suffers from an undercount, that the undercount is differential across race, ethnicity, and age, but that there is diversity within these groups. There are substantial and statistically different pictures of the population that are drawn by these three methods even at the national level. This is worrisome in and of itself. An adjustment based on the PES will be at face-value substantially different from our demographic estimates at the most aggregate levels. Whether it is an improvement depends not on its ability to add people and to subtract people from the census, but, rather, on its ability to add them and subtract them from the right places.

The Quality of the Census Enumeration

Special Advisory Panel Members Ericksen, Estrada, Tukey and Wolter all condemn the census as being fatally flawed.[17] I concede the census' imperfections, but the critical inquiry under this guideline is not how flawed the census is, but whether the PES can fix it.[18] Census taking is a complex task that must be completed within a short period of time. In an operation employing 350,000 temporary workers spread over more than 400 offices across the country, quality control is a real problem. The management information system the Census Bureau installed allowed the Census Bureau and the panelists to have access to the type of data panelists report. Thus, while identifying the flaws in the census is important for planning the next one, it simply begs the question that Guideline One poses: Is there convincing evidence showing that the adjustment is more accurate than the enumeration?

The Quality of the Alternative Measurement Tools

In considering whether to adjust the population for undercounts, the quality of the tools used to measure and then make an adjustment is important. The two methods that are alternative to the census are DA and the PES.

Demographic Analysis

Demographic analysis is a count of the aggregate population that is not based directly on any census. Instead it is built from administrative records including birth and death certificates, immigration records, and medicare records, among others. Limitations in record-keeping limit demographic breakdowns to those by age, sex, and black/non-black. There is no uniform reporting of ethnicity (*e.g.*, Hispanic origin) or the race of children of biracial couples. Even the same person might be reported as having different characteristics on birth and death records. Because we do not keep records of movements of individuals within the United States this analysis can only be done at the national level.

Furthermore, demographic estimates of the population are continually being changed. No demographic estimates are ever final, as new sources of data and statistical models are used presumably to improve the inferences made about the population. (For example, as a result of the demographic analysis for this census, the estimates of the 1980 population were still being changed as late as last month.) This year the Census Bureau undertook a series of investigations into the quality of the demographic estimates. An important improvement in the estimates was the first attempt to characterize the uncertainty inherent in them with uncertainty intervals about the point estimates. These improvements are reported in the demographic reports D1–D11.[19] Because demographic analysis will not be used in any adjustment, any detailed discussion of its results is foregone. Nevertheless, it is worth noting that the uncertainty intervals have been used in the previous descriptions of consistency of the various estimates of the population.

In an article in *Science*, David Freedman, Professor of Statistics at the University of California at Berkeley, discusses the limitations of DA in some detail.[20] Racial classification procedures vary widely. Incomplete coverage of vital statistics is a problem especially for certain age groups, with further variation by race and sex. In fact the census is used to adjust the birth certificate data that go into DA before DA is used to evaluate the census. Wachter also notes the complexity of DA,[21] and the fact that it is rightly subject to continual revision. He is particularly uncertain about the correctness of the estimates of immigration. He applauds the innovations in the 1990 DA, but quotes his colleague, Wolter, as saying: "The corrections that have been made are indicative of the corrections yet to be named."

The Post-Enumeration Survey and Dual-System Estimates

The Post-Enumeration Survey serves two related purposes. It is used as a measure of the accuracy of the census and it is used together with the census and statistical methods to generate adjusted counts. These adjusted counts are technically referred to as dual-system estimates (DSE). To evaluate the quality of the PES a series of 21 studies was done.[22] There are two questions that the Census Bureau intended to answer with the evaluation of the quality of the PES. First, whether the survey itself was of high enough quality in design and operation to be able to tell us something reliably about the faults of the census. Second, whether the adjusted counts or DSE were significantly more accurate than the census.

The Quality of the PES Survey

The 21 Census Bureau studies were designed to address the issues of quality in the PES and the DSE, some in a quantitative way, some in a qualitative way. They generated volumes of data that have not yet been fully analyzed or understood. Nevertheless, they have generated the basic material on which a judgment must be made regarding a possible adjustment of the census and the effect of that adjustment on the accuracy of the census.[23] In addition some of the panelists did their own studies on various aspects of PES quality. The broad picture that emerges from the analysis of these studies is that the PES was a generally high-quality survey that was well-executed.[24] There is little doubt that the PES detected an overall undercount in the census and a differential undercount at the national level by race and ethnic origin. But there are some problematic areas and disagreements among experts inside and outside the Census Bureau that have an impact on assessing the quality of the adjusted counts generated from the PES.

[17] Ericksen, page 2; Estrada, page 2; See also Appendix 3: Tukey, John W. "Recommendation on 1990 Census Adjustment," Member, Special Advisory Panel, June 18, 1991, (hereafter Tukey), page 3; Ericksen, *et al.*, pages 4–6.

[18] Nevertheless, this was at least the second-best census ever conducted.

[19] See the executive summary of these evaluation projects in appendix 2.

[20] See Appendix 13: Freedman, David A. "Adjusting the 1990 Census," Science, Volume 252, May 31, 1991, [hereafter Freedman] pp. 1233–1236.

[21] Wachter, pages 14–16.

[22] See the executive summary of these evaluation projects in appendix 2.

[23] Under Guideline Six, as explained later, "[i]f sufficient data and analysis of the data are not available in time to publish adjusted counts by July 15, 1991, a determination will be made not to adjust the 1990 census."

[24] See for example Ericksen, *et al.*, pages 13–16 for a good summary of the merits of the PES as a survey. Also see Wachter, page 2.

Missing data. The PES generates its estimate of the undercount by trying to ...ch households it has information ...l to households in the census. A ...ehold survey in the PES that is matched to the census record of that residence means there was no undercount of that household. A non-match means there was an undercount. Matching is a difficult process and sometimes it is unclear whether there is a match or not. It is not an automatic process, rather it requires judgment and discretion. (For example, is a household headed by R. Smith the same as one headed by Bob Smythe?) Ideally, each household in the PES is matched or adequately resolved as not matched and thus missed in the census. Any case which is not resolved becomes "missing data" and, thus, whether those cases would add to or subtract from the undercount is unknown. The lower the missing data rate is, the more accurate the results are presumed to be. Three evaluation projects examined this problem.[85] In general, missing data were not found to be a serious problem; [86] however, there were two troubling findings. First, it is standard practice to impute persons into unresolved match households. The imputation rates for the two parts of the PES, called the "P" and "E" samples, were high: 1.7% and 2.1% respectively, ...ch is equivalent to 3,900,000 and ...5,000 individuals in the census when ...eighted up to the national population total estimate. These numbers are the same order of magnitude as the undercounts. Second, the percent of imputation in an evaluation stratum is highly correlated with the size of the undercount in that stratum. Thus, the strata for which there is more doubt about the quality of the adjusted data because of imputation tend to be the same strata for which an adjustment would result in large increases in the population.

Although Ericksen, Estrada, Tukey and Wolter do not find missing data or imputation to be a problem, Wachter raises some basic questions about imputation.[87] The imputation scheme used for the PES is based on a series of assumptions that are mostly guesswork.

Given the assumptions, Wachter finds that this work is of high quality, yet he is hesitant to believe that these assumptions are necessarily valid. To get some idea of whether the assumptions are important he calculates strict upper and lower bounds on the effects of imputation.[88] This analysis shows that if the imputation assumptions were incorrect, the variation in the estimates could be well beyond that expected from sampling error alone. Thus these untested assumptions are critical. They may in fact be correct, but if they are not, the adjusted estimates may be significantly in error. This implies that the estimates in the adjusted counts are subject to more potential error than has been computed.[89]

Matching error. Highly accurate matching is important because matching errors in even a small percentage of cases can significantly affect undercount estimates.[90] Ericksen, Estrada, Tukey and Wolter find the matching process to be of high quality.[91] Although Wachter does not dispute that this is what the studies show, he believes that the estimate for the matching error is too low, because the rematch study "does not, by its nature, expose certain inevitable kinds of matching errors." [92] For example, he notes that the structured nature of the PES interviews could lead to inaccurate and inflated estimates for undercount rates. His evidence, though anecdotal, is suggestive of the fact that the variance due to matching error is conservatively estimated in the total error model.

Erroneous Enumerations in the Census. Erroneous enumerations include people who died before or were born after census day, fictitious people and pets listed as members of a household, twice counted people as well as people enumerated outside the PES matching area. There were a large number of erroneous enumerations in this census and they were differentially distributed. "While the national rate of erroneous enumerations was estimated at 5.4 percent, the rate for Blacks, Hispanics and central city Asians was 7.7 percent compared to 4.4 percent for all others. Minorities in central cities had the highest erroneous enumeration rate at 8.4 percent." [93] The Census Bureau studies indicate that the PES was good at detecting erroneous enumerations, although three processing offices show statistically significant underestimates of erroneous enumerations.[94] The national effect of these errors is small, but the impact on regional totals is unknown.

Ericksen, Estrada, Tukey, and Wolter take the large number of erroneous enumerations as an indictment of the census.[95] Although it is certainly a matter of concern, especially for future census planning, the relevant question is whether the large numbers of erroneous enumerations would affect the accuracy of the proposed adjustment. Wachter considers this question at length.[96]

> Erroneous enumerations and cases with insufficient information are not part of the usual statistical framework for dual-system estimation. Their modeling has received much less attention than the omission rates . . . The PES, however, turns out to show that erroneous enumerations account for a large portion of the variations in net undercounts across areas and post-strata. *This outcome very much complicates the task of understanding and assessing the adjustment process.*[97] (emphasis in the original)

The adjustment factor for a post-stratum is determined by the netting out of two kinds of errors in the census—in technical terms, gross omissions minus erroneous enumerations. One would hope that the predominant determinant of the adjustment would be the number of people missed in the census: areas with high miss rates get high adjustments. What Wachter demonstrates is that the erroneous enumerations—the number of extra people counted—are what is really driving the adjustment: areas with low duplication rates get high adjustments. For example, the three regions with the highest omission rates have very different adjustment rates. Like Wachter, I find it disturbing that "erroneous enumerations account for a large portion of the variations in net undercounts across areas and post-strata." [98]

As Wachter notes, Ericksen, Estrada, Tukey, and Wolter take the high levels of erroneous enumerations as evidence that coverage improvement programs were not finding real people but just adding fictional people to the count.[99]

---

[85] See executive summaries of P1, P2, and P3 in Appendix 2.

[8]6 Estrada, pages 11–13.

[87] Wachter, pages 21–22.

[88] Wachter, page 22.

[89] In a letter submitted on July 11, 1991, Ericksen and Tukey dispute Wachter's concerns over imputation. Professor Wachter was offered an opportunity to respond in the interest of fair play. In his rebuttal letter, submitted on July 12, 1991, Wachter stands by his statements. Both letters are contained in Appendix 16. Wachter correctly notes that his claim was only that "a great deal rests on the correctness of the assumptions in the imputation," not that his alternatives were more reasonable than the ones used.

[90] Comments by Barbara Bailar, Journal of American Statistical Association, (March 18, 1988). Pages 109–123.

[91] Ericksen, page 13.

[92] Wachter, page 20.

[93] Ericksen, et al., page 8.

[94] Estrada, pages 16–17; and the executive summaries of the evaluation studies P9 and P9a in Appendix 2.

[95] Ericksen, et al., pages 7–8.

[96] Wachter, pages 13–14.

[97] Wachter, page 11.

[98] Wachter, page 11.

[99] Ericksen, et al., pages 5–6 and Wachter page 12.

Wachter finds very mixed evidence on this question in comparing the counts in Detroit and Chicago. Late in the census enumeration, Detroit mounted an intense campaign to improve coverage, exceeding that mounted in Chicago. In the aggregate, Detroit did have a slightly higher erroneous enumeration rate, but a much lower omission rate. Thus, coverage improvement may very well have worked. However, for some categories of people, omission rates are roughly the same between the two cities, whereas erroneous enumeration rates are not. Thus, the evidence about coverage improvement is certainly more mixed than Ericksen, Estrada, Tukey and Wolter claim.[40] It is worth noting that Detroit and Chicago are lumped together when adjustment factors are calculated, despite their sizable differences in coverage patterns.

Correlation Bias. To the extent that the PES misses the same people that the census misses it will underestimate the undercount. The technical term for this problem is correlation bias. There are several ways of assessing the extent of this problem, but the basic message given by all of them is the same. There is strong correlation bias in the PES, especially among black males.[41] Ericksen, Estrada, Tukey and Wolter tend to dismiss this problem by noting that the presence of correlation bias results in an underestimate of the undercount, so an adjustment at least goes part way toward solving the problem.[42]

However, the presence of large correlation bias poses a fundamental difficulty for the adjustment procedure. Since there is no way to observe these people directly, the adjustment estimator attempts to include an estimate of these people. They are often referred to as the "4th cell" since they appear in the 4th cell of a 2 by 2 table in which persons in a particular post-stratum are classified as being in or not in the census and in or not in the PES. Unfortunately we have no direct data to verify if the assumptions for estimating the 4th cell are met. One piece of data indicates there may be a problem we do not fully understand. Traditional wisdom has it that males are generally more subject to correlation bias, since past data support the observation that males are more likely to be missed in both the census and the PES.[43] But, in 1990, about one-half of the people added to the estimate of the population from the 4th cell are women. Thus there is reason to doubt that the "fourth cell" numbers are correct. If that were the case the accuracy of the adjustment would be indirect.

One also expects that the number of people added to the adjusted population from the 4th cell should be small and that the estimate of the total population should be "lower than the truth." This is because no one expects that the estimate to fully reflect people missed in both the census and the PES. In past censuses, that has been the case. However, for 1990, the data are not consistent with past experience. Almost 5 million people were added to the estimate of the total population from the 4th cell, and the PES estimate of the total population exceeded the estimate from DA—a very unexpected finding.[44] Taken together, these findings indicate there may be problems in the adjusted count estimates that are not fully understood.

Wachter devotes several pages to the issue of correlation bias or as he calls it "catchability error."[45] His technical worry is that the allocation of this error to the model that measures the total error in the PES is done in an arbitrary fashion. Specifically, the national totals for black and for non-black males in six age groups estimated from DA are divided by the corresponding totals for females. Under the assumption of no correlation bias for females, these ratios are then multiplied by the national totals from the adjustment estimate for females in each group to give the predicted total for males. The differences between these predicted totals and the totals for men given by the calculated adjustment are the resulting national estimates of unreached persons. The method assumes all unreached people are men. This allocation, which critically affects conclusions about the accuracy of the census, is not based on empirical evidence on the distribution of those persons not reached by either the census or the PES, but rather on a formula of convenience. There is no unique way of choosing an allocation scheme. The one chosen is not obviously bad, but whether it is good is speculative and has no basis in fact. Furthermore, the variation in the PES estimates contributed by correlation bias is computed for sex ratios in an "ingenious" but ultimately untenable fashion.[46] It uses the capture probability of those reachable by the PES and census to infer a capture probability for people who intend to evade both the census and the PES.[47]

Wachter's argument over this technical point takes him back to a more fundamental point raised earlier, and also raised by Special Advisory Panel Members Kruskal and McGehee.[48] The PES is based on a statistical technique called "capture-recapture" which is often applied to estimating wildlife, particularly the number of fish in a pond. Fish are caught, tagged, thrown back and some are recaught in a second catch. An estimate of the population of fish can be made from the number of fish who are tagged on the second catch. The analogy made for the adjustment mechanism is that the census is the first catch and the PES the second. The analogy is not close, and it is not routine to adapt the wildlife model to counting the population.[49] The problem that

---

[40] Wachter, pages 12–13.

[41] See the discussion above.

[42] For example, see Estrada, page 14.

[43] See the discussion of hard to count groups in C.F. Citro and M. L. Cohen, eds., The Bicentennial Census, National Academy Press, 1985, Chapter 5, especially pages 177–186, and pages 224–237.

[44] See also the earlier discussion regarding the differences between DA and the PES at the national level.

[45] Wachter, page 18.

[46] Wachter, pages 18–19.

[47] In their letter submitted on July 11, 1991, Ericksen and Tukey dispute Wachter's concerns over the consistency of DA and the PES. In his rebuttal letter, submitted on July 12, 1991, Wachter stands by his statements. Both letters are contained in Appendix 16. It is difficult to referee this dispute at the eleventh hour, especially since the lateness of the Ericksen/Tukey letter gave little chance for Wachter to prepare a detailed response. It seems however, that even given the recognized inability of the PES to reach certain black males, a PES-based adjustment would have more persons than demographic analysis would indicate. Now suppose, in fact, that one were to use the behavior of those captured by the PES to extrapolate to those missed by both surveys, as Ericksen and Tukey suggest. The estimate of the population would be, at least by Wachter's estimate, yet another half-a-million higher. Then the PES would exceed DA by well over a million people.

Ericksen and Tukey also take Wachter to task for asserting that "[t]here is no evidence we know of that indicates that a substantial proportion of those persons counted neither by the PES nor the census avoided being counted." Ericksen and Tukey have apparently overlooked a well known study by Valentine and Valentine that concludes "one cannot [always] expect traditional interview or self-enumeration procedures to identify individuals of the type missed in the study area. * * * [T]he men were not reported because identification * * * could be detrimental to the economic welfare of the household." Citro and Cohen, *op cit.* pages 236–37.

[48] See Appendix 2: Kruskal, William, "Recommendation to the Secretary on the Issue of Adjusting the 1990 Census," Member, Special Advisory Panel, June 13, 1991, [hereafter Kruskal], page 2; Wachter pages 18–20; and also Appendix 8, McGehee, J. Michael, "Report to Secretary Robert A. Mosbacher on the Issue of Adjusting the 1990 Census," Member, Special Advisory Panel, June 21, 1991, [hereafter McGehee], pages 6–12.

[49] Citro and Cohen, *op cit.*, page 147, make this point clearly.

worries both Wachter and Kruskal is ...t, using the fishing analogy, some fish arder to net than others.[50] There among fish, some "wily trout" which cannot be caught at all. Similarly some persons are harder to count than others, and some impossible.[51] For a variety of reasons they avoid the census and other forms of registration. The conclusions drawn about the population depend on what assumptions are made about these unreachable people. Different assumptions lead to widely differing results.

McGehee's concern about the application of capture-recapture is related to this notion of countability. The census and enumeration are both done by enumerators of varying skills, in different kinds of geographical areas (urban, rural, inner city, suburb) in an attempt to enumerate people who have different incentives to cooperate with the census or the PES. Thus there is inherent in the process a large variation in the probability of a particular person being enumerated in a particular place by a particular census worker. Further, to see if a person was counted both in the census and the PES a match has to be made—we do not tag people like we tag trout.[52] The ability of the matcher thus comes into play here. McGehee recognizes that there are elaborate anisms in place to control for all potential variation, but many of those mechanisms depend on unverified statistical assumptions about what is important, and are changed after the data are in or after new research is completed.[53]

Total Error Model. An effort was made to produce estimates of expected error in the PES and variability of the estimates derived from the PES in project P16. This is generally referred to as the *total error model* since it was an attempt to combine the errors found in the PES by the other evaluation studies. These estimates of error cannot be made for any detailed groups. Instead, the population is divided into thirteen very broad categories called evaluation strata.[54] The estimates of errors for each evaluation strata are meant to be indicative of the uncertainties due to sampling error and all known components of non-sampling error. Whether the results of this study of large groups holds for smaller groups such as post-strata, states, cities or districts is unclear.[55]

This evaluation technique represents pioneering work on the part of the Census Bureau. It has been refined several times since the beginning of June, and every indication is that more refinements will be made as research on it is completed over the next several months. Nonetheless, some conclusions can be drawn from this project. On the one hand, the errors introduced by measured flaws in the PES process seem small. On the other hand, the model does show that the PES is biased toward overestimating the undercount and that a bias-corrected estimate of the undercount would be about 1.4 percent rather than the production estimate of 2.1 percent. This means about a third of the net undercount adjustment in the DSE comes from bias in the PES.

Furthermore, the undercounts tend to be higher in the minority evaluation strata, as are the biases in the PES. Even after bias correction, the minority evaluation strata show statistically significant undercounts. Ericksen, Estrada, Tukey, and Wolter note that the shift in shares of each evaluation post-strata would be small if the production estimate were corrected for bias.[56] Wachter [57] expresses various concerns about the computation of the total error model and its components as does the minority of the Undercount Steering Committee.[58] The results of this model are used further in assessing the quality of the counts themselves.

The Quality of the Adjusted Counts

The fact that the PES was generally a high quality survey does not necessarily imply that it results in high quality adjusted counts. To the contrary, erroneous enumerations and correlation bias lead to the conclusion that there are serious doubts about the quality of the adjusted population estimates.

To understand the statistical issues involved in assessing the quality of the adjusted counts it is necessary to begin with a summary understanding of three measures of the population that the Census Bureau compared.[59] First there is the census enumeration. Second there are the adjusted counts or the production dual-system estimates (production DSE). Third there is an alternative DSE that corrects for biases found in the production DSE by examination of the evaluation of the PES in the P-studies. The third measure is used to judge the relative accuracy of the census and the production DSE. There are two main elements of concern: (1) whether to test the accuracy of population totals or of population distributions and (2) how such tests should be performed.

Should population totals or population distributions be compared? Acceptance of the PES measure of the national undercount as reasonable is only a necessary—not a sufficient—condition for it to be an adequate instrument to be used to adjust the actual enumerations. There has always been an undercount in the census. The central questions for the Constitutional and statutory purposes of the census are whether the undercount is evenly or differentially distributed across geographical areas and jurisdictions, and whether we know how to reduce the range of any differential undercounts. Indeed Congress has recognized this problem as well.[60]

These questions have not been squarely faced. For the most part, Census Bureau analysts concentrated on whether we know enough to reduce the errors in the numeric counts without regard to whether this increases or decreases the severity of differential undercounts across geographical areas or jurisdictions. That is, they interpreted accuracy as concerned with getting the number of people closer to the truth rather than getting the allocation of the population for the purposes of political representation and funding closer to the truth. The two do not necessarily go together.

An illustration of the problem with using the absolute criterion alone is useful. Suppose you observed an enumeration which missed exactly 5 percent of the people in each and every block. Thus, although 5 percent is missed in each and every block, the proportion of the total population in each block is still estimated correctly. Suppose now that you adjusted this enumeration by increasing the counts in half the blocks by 1 percent and increasing the counts in the other half by 5 percent. On average you would have reduced the undercount of the

---

[50] Kruskal, page 3; and Wachter, page 18.

[51] See Citro and Cohen, *op cit.*, pages 130–142.

[52] Although often trout lose their tags which poses a similar conceptual problem.

[53] McGehee, pages 8–12.

A list of evaluation strata and their component strata are included in the Decennial Census edural Documentation, below.

[55] Wachter, page 18.

[56] Ericksen, *et al.*, page 15.

[57] Wachter, page 17.

[58] Undercount Steering Committee, page 8.

[59] These measures will be explained more fully in the course of this discussion. The alternative DSE is also called the "target" population in Census Bureau documents.

[60] Subcommittee Chairman Thomas Sawyer, for example, noted that "If the undercount were evenly distributed geographically and demographically across the population, it probably would not pose the problem that we confront here and the difficulty that we face in asking the Secretary to come to this decision." Hearing before the Subcommittee on Census and Population of the Committee on Post Office and Civil Service, U.S. House of Representatives, January 30, 1990. Serial no. 101–43, page 18.

population by 3 percentage points thus, improving the numeric accuracy of the nationwide total. The numeric accuracy of the absolute level of the count also would have improved for each block. However, the block proportions would now be wrong. Half the blocks would be 2 percent too small and half would be 2 percent too large relative to the average undercount. The absolute criterion would prefer this type of adjustment even though it moves from a situation in which every citizen gets his or her fair share of representation and funding to one in which every citizen got 2 percent too little or 2 percent too much.[61]

It is quite possible this kind of error could occur when the PES misses persons. The PES failure to include large numbers of black males in the adjusted counts could have caused just this kind of error. We simply do not know if it did.

I conclude that the Constitutional and legal purposes for the census must take precedence, and accuracy should be defined predominately in terms of getting the proportional distribution of the population right among geographical and political units. This argues for putting aside the judgment of accuracy based on getting absolute numbers right (numeric accuracy) and instead focusing on the question of whether there is convincing evidence that the accuracy of the population distribution in the adjusted numbers (distributive accuracy) is superior to the distributive accuracy of the actual enumeration. The quality of the adjusted counts themselves must be examined to address this important issue squarely.

What is the criterion for accuracy? Guideline One mandates that the census enumeration "shall be considered the most accurate count of the population of the United States, at the national, State, and local level, unless an adjustment is shown to be more accurate." This guideline requires a series of statistical hypothesis tests at various levels of geography in which the adjusted counts are to be presumed less accurate measures of the population than the actual census enumeration unless there is convincing evidence that the adjusted counts are closer to the true counts than the actual enumeration.

The true population counts cannot be observed. However, classical statistics provides a standard way of approaching the required inference. In accordance with Guideline One, we take as a working (null) hypothesis that the actual enumerations in fact better characterize the true population. The adjusted counts are an alternative measure and the question is whether the available evidence permits us to reject the hypothesis that the census better describes the true population.

We shall see below that the Census Bureau has provided substantial (although not necessarily "convincing") evidence that the adjusted counts are more accurate if accuracy is interpreted to mean numeric accuracy. However, the evidence provided by the Census Bureau tends to support the superior distributive accuracy of the actual enumeration. Thus, since accuracy is interpreted in terms of the fairness of the implied distribution of representation and funds, the Census Bureau report supports the conclusion that the adjusted counts are not more accurate.

The choice of accuracy criterion is crucial because there appears to be a substantial national net undercount in the numeric census counts. Simply correcting for the estimated net undercount can improve numeric accuracy but significantly worsen distributive accuracy. We can see that we missed people in most areas, but we lack a tool which can improve the distribution of population for the purposes of political representation and funding.

How are the tests of accuracy performed?

(a) The Census Bureau Loss Functions

The Census Bureau approach to testing the quality of the adjusted counts relies heavily on showing that the PES was well-executed and that the identified biases in the production Dual System Estimates or adjusted counts (DSE) are small relative to an "ideal" DSE. Unfortunately this type of validation methodology does not work in the present instance because of a basic design flaw: The DSE fits broadly into the class of "certainty-equivalent" predictors which use estimates as if they were known for certain rather than subject to statistical variation. A statistically optimal estimate of the population for an area would take account of this uncertainty.[62] Thus the conclusion that the measured shortcomings of the adjusted counts under consideration (the "production DSE") are small relative to the ideal DSE merely means that the production DSE has a chance of improving accuracy. It is unacceptable to go the next step and conclude that a good production DSE would be more accurate than the actual enumeration.

The production DSE are in fact less accurate than those ideal DSE because (a) the data were less than perfect, and (b) the correct model was not known. The bulk of the Census Bureau effort was aimed at seeing whether these data and modelling problems were disqualifying for the production DSE. It is clear that the production DSE are not unbiased estimates of the differential undercount rates and the DSE procedure overcorrects for the measured undercounts. This is measured in the total error model discussed above. These biases are quantified for thirteen large evaluation strata.

Using the total error quantification, the Census Bureau has generated an alternative Dual System Estimator of the population. It is worth noting here, that the errors in the production DSE are quantified for 13 very large groups of people. These errors are then "parcelled out" to the 1392 post-strata used to calculate an adjustment, the adjustment factors are corrected for these biases, and the alternative DSE is calculated. Since there are also estimates of variance for the DSE, the Bureau actually calculates a statistical distribution of possible alternative DSE. A thousand random draws from this alternative distribution were used to generate estimates which the Census documentation terms "the target population." This is not the true population distribution—which is unobservable—but rather a tool for assessing the quality of production DSE counts relative to an "ideal" DSE based

[61] Kruskal gives a similar example on page 7 of his recommendation. See also Citro and Cohen, *op cit.*, page 318. "While synthetic estimation is suggested for adjustment, because of its arithmetic and computational simplicity, synthetic estimation is not necessarily an improvement over the census count." Cohen and Citro use a numerical example as an illustration.

[62] The optimal estimate would average the ideal DSE estimate (based on the correct model and perfect data) with the actual enumeration with more weight being put on the actual enumeration when the model parameters are less precisely estimated. In point of fact, there are statistical theorems which demonstrate that even if the correct statistical model were known and perfect data were used, the Dual System Estimator (DSE) could generate adjusted counts which are either (1) clearly less accurate, or (2) not significantly more nor less accurate, or (3) clearly more accurate measures of the true population than the actual enumeration from the census. The question is which of these occurred? A textbook analysis of the suboptimality of the certainty equivalent approach is found in Arnold Zellner, An Introduction to Bayesian Inference in Econometrics, New York: John Wiley & Sons, 1971, pages 322–327. Intuitively, the problem arises because a full correction is attempted which is optimal only if one knows exactly the undercount or overcount in each area. As the actual uncertainty increases about exactly where and how many people were missed, attempts to make the full estimated correction increase the error variance relative to optimal and eventually, if uncertainty is large, relative to the unadjusted counts.

on more perfectly measured data and more correct models. But this hypothetical DSE is also just an estimator—subject to statistical error. So a correct analysis must account for two errors: (1) the error that comes from using the production DSE rather than the idealized DSE and (2) the error that is inherent in the idealized DSE. Then that combined error should be compared with the error in the actual enumeration.

To make matters even more complicated, legislative—and, now, judicial—representation must be apportioned and allocated over many levels of government into districts that treat their residents as fairly as practicable. Thus, comparisons must be made not only at the various levels of government on which funding is based, but down to the census blocks which are the basis for drawing district lines for Federal, State, and local elections. Unfortunately, the Census Bureau did not have the time to conduct the hypothesis tests required by Guideline One before the Undercount Steering Committee report was completed on June 21, 1991. The method they used instead to make these comparisons is called loss function analysis.

In brief, loss function analysis is used to compare two sets of counts for the same population. Ideally, one of the sets of counts is the true population, and thus loss in accuracy from using the alternative set of counts is measured. In practice, however, the truth is not known, so care must be taken in the interpretation of results. A loss function analysis can be performed at any level of geography—states, counties, cities, precincts, or blocks.

As an example, suppose a loss function analysis is being calculated for states. The difference between the two estimates of the population is calculated for each state. Then some kind of average is taken of the differences across all states to get an aggregate measure of total loss. The differences may be squared, summed and the total divided by the number of states. Alternatively, the absolute values of the differences may be averaged, where the average is weighted by the size of the state. There are an infinite number of formulas that can be used to average the state-by-state losses to get a single measure of total loss. These formulas are called "loss functions," and the results of any analysis can depend heavily on which loss function is chosen. For example, the loss function that uses squared differences penalizes a few large errors much more heavily than many small errors. The absolute value loss function does not have this property. The choice of a loss function is not scientific. It is usually made on the basis of convenience or tradition.

One more general comment on loss function analysis is needed. The loss function is ideally suited to measuring loss when an estimator of a population count is being compared to a known true count. In this case, the interpretation of the loss is straightforward. It is the accuracy lost by using the estimator. However, when one imperfect estimator is being compared to another, it is more difficult to interpret the loss of one estimate. The temptation is to call one estimator the "truth" and measure loss against it. But one is not measuring loss against the truth. This is simply measuring loss of one estimate against another. There is no reason to think this analysis tells you anything about the truth. In loss function analysis, it is critical to consider the base being used for comparison—losses are measured only relative to that base.

The loss function analysis run by the Census Bureau asked whether the enumeration or the production DSE was closer to the "ideal" DSE.[63] This does not form a statistical test of whether the production DSE are more or less accurate than the census counts. It only calculates which set of numbers on average is closer to another set of estimates (the target population). These tests were simply not proper statistical tests to address the critical hypothesis about the distributive accuracy of the PES and the census enumeration.

Their examination of this closeness question erred further in two significant ways: (1) Instead of comparing the production DSE that would be used, they compared the mean of 1000 draws from a model reflecting the statistical properties of the DSE. This effectively eliminates the inaccuracies derived from using one particular set of adjustments. (2) Rather than using Guideline One's mandate that the actual enumeration be deemed more accurate unless the adjusted counts are shown convincingly to be more accurate, the Census Bureau did the reverse—they preferred adjusted counts if the actual enumeration was not proven more accurate.[64] Thus the Census Bureau loss function analysis was seriously deficient.

There is, nonetheless, a June 27, 1991, Addendum to the Undercount Steering Committee report of June 21, 1991, that corrects some initial flaws in the loss function analysis.[65] This addendum attempts to correct for the error in failing to allow for the fact that the target population was itself an estimator subject to random variance. An allowance for this variance was removed from the variance charged to the census counts and estimates made of the number of states for which the population proportion would be made less accurate was generated. The number of state proportions worsened depends crucially upon the allowance made for variance in the alternative DSE: If only the variance measured in the total error model is used, then the shares of an estimated 21 states are made worse by adjustment (using an absolute value loss function).[66] However, this is clearly a minimum estimate. "As a matter of judgment, the total understatement of variance of the estimates from the smoothing model may be in the range of a factor of 1.7 to 3.0 in terms of variance," according to the Undercount Steering Committee.[67] Allowing for a variance factor of 2.0, which is near the lower end of the Undercount Steering Committee range, the proportional shares of about 28 or 29 states would be worsened by an adjustment in terms of distributive accuracy.[68]

Even with the variance factor set at only 1.0, adjustment is estimated to have worsened distributive accuracy compared to the census counts in 11 of the 23 metropolitan areas in cities with 500,000 or more persons: Phoenix, Washington, DC, Jacksonville, Chicago, Baltimore, New York City, Memphis, Dallas, El Paso, Houston, and San Antonio. Again using only the measured variance, half of the 14 metro areas in counties with over 500,000 persons are made less accurate proportionally by

---

[63] This loss function analysis is described in detail in Undercount Steering Committee, pages 6–7; and Bryant, pages 12–14.

[64] This last error may reflect the fact that the Census Bureau ignored the difference between the true population and its own approximate ideal estimator. See for example, the Undercount Steering Committee, page 2: "Time did not allow for full simulations of accuracy for smaller areas. There was some evidence from the loss function analysis, but there was no independent evidence with which to compare it. . . . Even so, in the absence of direct evidence to the contrary, the majority concludes that adjusted counts are generally more accurate at lower levels."

[65] A discussion of how this change affected the Undercount Steering Committee's conclusions is contained in the discussion of Guideline Six, below

[66] See Appendix 5, Addendum to the Undercount Steering Committee Report, July, 1991 (hereafter Addendum), page 3. Given the original erroneous analysis, the Undercount Steering Committee report (page 6) was formulated when the committee thought the accuracy of only about 11 states was worsened by adjustment.

[67] Undercount Steering Committee, page 5. The actual variance is believed to substantially exceed the measured variance because of doubts similar to those raised by Wachter for the matching and imputation procedures.

[68] Addendum, page 4.

adjustment. Only aggregate measures are available for areas of other sizes. These show that on average the adjusted figures improve distributive accuracy relative to the census, but no detail is given as to the number of jurisdictions for which the PES is closer than the census. In all these sub-state cases, too, the estimated distributive accuracy of the adjusted figures deteriorates dramatically compared to the census if the variance is increased to allow for the unmeasured uncertainty in the estimator.

In sum, the corrected Census Bureau estimates of distributive accuracy marginally favor the adjusted counts—though many states and communities would be less accurate—*if* only the measured variance is used. When the variance is increased into the plausible range (in the professional judgment of the Undercount Steering Committee), distributive accuracy comparisons are more favorable to the census counts.

It is worth reiterating that Guideline One specifically places the burden of proof on the adjusted estimates, not on the census. The census is considered to be more accurate unless the adjusted figures are shown to be more accurate. With respect to places under 100,000 population there is no direct evidence that adjusted counts are more accurate.[69]

What evidence there was based its conclusions primarily on the numeric accuracy of the adjusted counts rather than the adjusted proportions, and that the Bureau depended upon indirect evidence rather than direct tests of statistical hypotheses.[70]

(b) Face validity tests

In addition to Loss Function Analysis computed by statisticians, demographers at the Census Bureau made an independent evaluation of the adjusted population counts for states. To do this they compared the adjusted state counts with counts simulated by DA. To make the simulations (because DA provides data only at the national level), they disaggregated census counts for each state by race and Hispanic ethnicity. They then applied DA national undercount rates to black and non-black subpopulations and PES rates to Hispanic and Asian and Pacific Islanders. Then they built up new state estimates by recombining the racial and ethnic groups. These simulated state estimates further confirmed the "face validity," or reasonableness, of the adjusted state counts.[71] These face validity tests depend critically on what the analyst expects. Face validity tests certainly cannot be a substitute for formal tests, but just as face validity can be used to show that adjustment is making counts more accurate, face validity can show the opposite.

For example, is it reasonable that New Mexico has the highest undercount rate of any state? Why should the undercount rate for Montana be higher than that of New York State? How can the very low estimated undercount rates in cities like Philadelphia be explained? Of the large cities, only Washington, DC and Boston showed increases in their black populations between 1980 and 1990. Yet, Washington DC is estimated to have a very large undercount rate and Boston is estimated to have a very small undercount rate. Why are the only estimated overcounts for cities over 100,000 concentrated in New England? Why should Akron and Dayton have high estimated undercount rates (3.0% and 3.3%, respectively) and Cleveland have such a low estimated undercount rate (1.4%)? These examples illustrate as above the point noted above that was raised by Wachter earlier—there is much more texture to the pattern of undercount that lies well beneath the surface of any aggregate loss function analysis. Face validity cuts both ways. And the face validity of the proportions of persons in states and localities has not even been checked.

(c) Ericksen, Estrada, Tukey, and Wolter's claims regarding accuracy

These panelists take a different approach to the problem of accuracy of the counts at state and local levels. An article by Wolter and Causey attached to their jointly authored document[72] argues that accuracy improves, on average, at lower levels, so long as the measured undercounts at aggregate levels tend to have smaller errors than the original enumeration. In addition it is argued in a similar manner in an attachment to the joint report that Ericksen, Estrada, Tukey and Wolter submitted that adjusted counts will on average improve block level data (and thus data for localities) consistent with its improvement of data at larger units of geography.[73] Thus their argument asserts that by applying the total error model to the 13 evaluation post-strata, the PES is shown to be more accurate than the census and the error in the PES is shown to be low. They conclude, based on the theoretical argument by Tukey and the empirical argument made by Wolter and Causey, that

a. The total combined error increases as the size of the group decreases; *e.g.*, the combined errors for 5 million blocks will be larger than the combined errors for 1392 post-strata.

b. Consequently, the improvement in amount due to adjustment will be nearly the same for larger and smaller groups, the improvement in percentage terms decreases, but does not change sign, as the groups become smaller.[74]

Ericksen, Estrada, Tukey, and Wolter note that these conclusions depend on a particular measure of combined error—a loss function that uses a size-weighted sum of relative error. Their primary point is that, with such an error measure, conclusions about local accuracy can in fact be drawn from accuracy conclusions at larger levels. In short, they contend, "improvement in quite large areas thus prophesies improvement in very small areas, as well as a variety of intermediate levels." They see a post-enumeration survey with small measured error (and some, like Wachter and the Undercount Steering Committee contend that such error is very conservatively measured) for thirteen large evaluation strata. They conclude that the adjusted counts for these large evaluation strata are more accurate—a questionable inference because they made no formal statistical test of this hypothesis. From this

[69] The Undercount Steering Committee report states "in the absence of direct evidence to the contrary, the majority concludes that adjusted counts are generally more accurate at lower levels." and later "while analysis was not available for smaller areas, the majority concludes that acceptable patterns would happen there also." (Undercount Steering Committee, page 2.). The reasoning is contrary to the explicit mandate of the guideline. Similarly the Director stated, "there is little evidence to judge whether the proportional distribution of adjusted counts is more accurate for places under 100,000. However, Loss Function Analysis shows that for metropolitan places of less than 25,000, 25,000–49,000 and 50,000 or more, and for nonmetropolitan places less than 25,000, and 25,000–49,000 in total, by these sizes categories, adjusted counts are more accurate than the census. However, there are concerns about the accuracy of the loss function assumptions for small areas." (Bryant, page 14.)

[70] In a June 28, 1991, memorandum Senior Mathematical Statistician Robert Fay reports his efforts at conducting formal hypothesis tests of the distributive accuracy of the adjusted figures at the state level only. There was not time for the Undercount Steering Committee to review this memorandum and it may contain further errors. Nonetheless, although the hypothesis tests rejected the superior distributive accuracy of the census counts if only the measured variance was changed to the adjusted figures, the superior accuracy of the census counts was easily accepted for a variance factor of 2.0 and appears (by interpolation) acceptable at any variance factor in the Undercount Steering Committee's plausible range of 1.7 to 3.0.

[71] Bryant, page 14.

[72] See appendix G of Ericksen, *et al.*

[73] See appendix F of Ericksen, *et al.*

[74] Ericksen, *et al.*, page 20.

questionable conclusion they apply mathematical theory to infer *average* accuracy improvements at lower levels.

In testimony before Congress, an official of the General Accounting Office raises some questions on the issue of sampling error and lower level geographic accuracy:

We believe the amount of sampling error, or variability, deserves attention by the Secretary because it was a consistently high source of uncertainty in the PES over- and undercount estimates. The PES estimates are based on samples and therefore subject to random error. The levels of sampling variation measured by the evaluations of the PES were generally much higher than anticipated by the original design of the PES. For example, even after smoothing to reduce sampling variability, PES over- and undercount estimates for 4 of the 13 evaluation groups did not show a statistically significant difference from the census count. In other words, due to the variability resulting from doing a sample, *the Secretary cannot be sure whether 4 of the 13 population groups reviewed in the Bureau's evaluation of total error in the PES were overcounted by the census, undercounted, or if the census count was correct.* (emphasis added)

The need for precision is especially important because the Bureau's procedure for carrying down PES adjustment factors to lower geographic levels applies the same adjustment factors to large numbers of people over wide geographic areas with similar demographic characteristics.[75]

The Wolter/Causey paper does not address this argument directly. In addition, Wachter argues cogently against indiscriminate use of the Wolter/Causey paper:

Theirs is a very interesting paper, but its relevance is limited by its concentration on highly aggregated summary measures of improvement. It does not present explicit results on how many units at various levels might be made worse and how many made better by an adjustment. Furthermore, important calculations in the paper depend on stylized assumptions about correlations in the components of the undercount which may or may not hold in fact either for previous PES-like data or for the 1990 PES. These prior studies are valuable, but they are no substitute for examination of the actual 1990 data.[76]

There are fundamental difficulties with the Wolter/Causey argument. I am not convinced that at the evaluation strata level we can conclude the PES is more accurate. First, the measured bias alone is one-third of the total undercount and the Undercount Steering Committee itself stated that there are other non-measured sources of error.[77] Wachter also raises several fundamental concerns about this measurement. Second, the analysis depends on a particular loss function that weights a few large relative errors more than many small ones. This is not inherently bad, just arbitrary.[78] Wachter perhaps summarizes it best:

I do not believe that any highly aggregated index or loss function is appropriate for summing up overall accuracy. It is informative to understand how much the outcomes of calculations with different versions of such aggregated indices differ. But the choice among them is not a scientific choice. Each such index involves implicit value judgments about different sorts of error. For example, each index determines whether a few large errors are more serious that a great many smaller errors. *Whether we agree with a particular tradeoff is a matter of personal and political values. It should not be disguised as science.*[79] (emphasis in original)

Loss functions mask the incredible complexity of the adjustment operation behind a single number. To get a glimpse of this complexity, it is useful to look at the undercount rates by state. Table 1 and Figure 2 of Appendix 10 show the undercount rate by state with margins of error. Counting the District of Columbia as a state, 42 of the 51 states show an undercount rate that is statistically significant. More importantly, however, is how these undercount rates differ from the national average, since it is these differences that determine which states win and which lose. Table 6 and Figure 1 of Appendix 10 show these differences again with margins of error. Only 18 of the 51 states have an undercount rate that is significantly different from the national average. That means in 33 states we do not know if the undercount rate is higher, lower or the same as the national average. Put another way, we do not know if these 33 states deserve more or less political representation and Federal funding than they are receiving. We do not know for these 33 states if an adjustment would result in a more equitable distribution of political representation and resources.

There are winners and losers from an adjustment—that is to be expected whenever a fixed set of resources is going to be divided. More seriously, however, there is general agreement that there will be some localities' counts that will be made less accurate by an adjustment. The proponents contend that, on average, more areas are made accurate, or more people live in areas whose counts are more accurate, or on average the counts are more accurate. These are all vague and general statements that do not describe the areas of the country where accuracy is likely increased and decreased, the types of towns where accuracy is likely increased and decreased, the neighborhoods where accuracy is likely increased and decreased. We have already seen above that general statements about improved accuracy on average are little if at all justified if realistic values are used for the error variance of the alternative DSE. Furthermore, the rhetoric, if not always the analysis, is centered around absolute levels of the counts, not improvements in the distribution of the counts.

Conclusions

Guideline One requires that convincing evidence be offered that the adjusted estimates of the population are more accurate than the census at the national, State, and local levels. In the absence of such evidence, the census counts are concluded to be the most accurate.

At the national level, it is likely that the PES-adjusted estimates reflect more accurately the total population and the racial and ethnic populations of the country. It appears equally clear, however, that the PES omitted large numbers of certain groups—notably black males. We have no information on the location of these persons. In addition, the PES and demographic analysis lead to sharply different conclusions about the accuracy of the census for several age/sex groups at the national level. Although these are not definitive disqualifiers at the national level, they do raise some question as to whether the adjusted figures are more accurate than the census count even at the national level.

The Constitution requires a census every 10 years not just to count the total number of people in the United States but to locate them so that political representation can be allocated to the states and the people in them in proportion to their numbers. I conclude that the primary criterion for accuracy should be distributive accuracy—that is, getting most nearly correct the proportions of people in different areas. Improved numeric accuracy, although in itself desirable, cannot compensate for treating states and individuals less fairly.

---

[75] See appendix 17. General Accounting Office. "1990 Census: Applying PES Results and Evaluations to the Adjustment Decision." Testimony before the Subcommittee on Census and Population. Committee on Post Office and Civil Service House of Representatives. [hereafter GAO Report]. Pages 7–8.

[76] Wachter, page 2.

[77] Undercount Steering Committee, page 8.

[78] Indeed, Erickson, Estrada, Tukey, and Wolter make no claim of uniqueness for their choice of loss function. As noted earlier, the choice of loss function can control the results of an evaluation.

[79] Wachter, page 5.

At the State and local level the correct statistical analysis for both distributive and numeric accuracy simply has not been completed. The total error model indicates that the adjusted figures tend to be too high but generally closer in numeric terms to the true population than the census counts which tend to be too low. However, there is sufficient uncertainty about the true variance of the adjusted figures that even numeric accuracy has not been definitively demonstrated. The loss function analysis and hypothesis tests that have been prepared by the Census Bureau to date, although of uncertain reliability, do support the superior accuracy of the census counts versus the adjusted figures when we consider distributive accuracy—or fairness—and use reasonable estimates of the error variance of the alternative DSE. That is, for the Constitutional purposes of the census the available evidence is consistent with the census counts being more accurate than the adjusted counts. There is certainly not sufficient evidence to reject the distributive accuracy of the census counts in favor of the adjusted counts.

I conclude that, in accordance with Guideline One, the census counts are the most accurate count of the population of the United States at the State and local levels. While the preponderance of the evidence leads me to believe that the total population at the national level falls between the census counts and the adjusted figures, that conclusion is not relevant to the determination of distributive accuracy. Thus this guideline weighs in favor of a decision not to adjust.

Guideline Two

The 1990 Census may be adjusted if the adjusted counts are consistent and complete across all jurisdictional levels: national, State, local and census block. The resulting counts must be of sufficient quality and level of detail to be usable for Congressional reapportionment and legislative redistricting, and for all other purposes and at all levels for which census counts are published.

Explanation

This guideline acknowledges that the population counts must be usable for all purposes for which the Census Bureau publishes data. The guideline also reinforces the fact that there can be, for the population at all geographic levels at any one point in time, only *one* set of official government population figures.

Thus, the level of detail must be adequate to produce counts for all such purposes. If the 1990 Census count is to be adjusted, it must be adjusted down to the census block level. It must be arithmetically consistent to eliminate confusion, and to prevent any efforts to choose among alternative sets of numbers to suit a particular purpose.

If the Census is to be adjusted, a process called synthetic adjustment will be used. A synthetic adjustment assumes that the probability of being missed by the census is constant for each person within an age, race, Hispanic origin, sex, and tenure category in a geographical area. A synthetic adjustment is performed in two steps. First, the preferred adjustment factors are estimated for a variety of post strata defined by age, race, Hispanic origin, sex, and tenure within geographic areas. Then the adjusted estimate in each category for a census block is obtained by multiplying the unadjusted census estimate in that category by the adjustment factor. The adjusted census estimate for the census block is computed by adding the estimated adjustments for each post strata cell of the block. Put simply, in an adjusted population count each individual enumerated will receive a relative weight according to his or her race, age, sex, ethnic background, tenure, and place of residence. The aggregate counts will then be built up from the weighted individuals to census block, local area, state and national counts.

We will conduct evaluations of small area estimations to ensure that this process results in counts that are in fact more accurate.

Evaluations of small area estimation. Coverage error may vary substantially within the post-enumeration-survey post-strata, although the post-strata were drawn to be homogeneous with respect to expected coverage error. The goal of this analysis is to determine whether or not the assumptions underlying a synthetic adjustment of the census are valid and produce counts which are more accurate at all geographic levels at which census data are used. In particular, the within-strata block-to-block variance in characteristics and net overcounts or net undercounts will be analyzed.

Discussion

If I had determined that an adjustment should have been undertaken, the Census Bureau would have issued block-level Public Law 94–171 tapes that would have replaced those issued in the first three months of this year. Replacement Summary Tape File (STF) data would have also been issued and all future census products would have used adjusted counts. Our ability to have done so would have satisfied the production requirements of this guideline.

The substantive question here is whether the adjusted counts are of sufficient quality to be used for all purposes for which census counts are published. Clearly the quality of the adjusted figures is intimately related to their accuracy, which, as the discussion of the preceding Guideline shows, does not compare favorably with the actual enumeration. This Guideline raises another issue—synthetic adjustment.

As explained earlier, the adjustment process uses a survey of persons in 5,290 block clusters to change the number of people in 4,830,514 blocks. Based on extrapolation from this survey 6,188,204 unidentified persons are added by duplicating records of people counted in the census, and 918,937 people who were actually counted in the census are deleted. The adjustment process is done by dividing the population of the country into 1392 groups. Each member of one of these groups is assumed to have the same probability of being missed in the census as every other member of that group. The real quality of the census in a given block or even a given city has little impact on the adjustment of the count of the population of that block or that city. As will be seen in the discussions of Guidelines Seven and Eight, most local officials think that the adjustment will fix particular problems that they have identified in the count for their towns. It would do no such thing.

A synthetic adjustment assumes that the probability of being missed by the census is constant for each person within an age, race, Hispanic origin, sex, and tenure category in a geographic area. These groupings of persons are called post-strata. A synthetic adjustment is performed in two steps. First, the preferred adjustment factors are estimated for 1392 post-strata. Then the adjusted estimate in each category for a census block is obtained by multiplying the unadjusted census count in that category by the adjustment factor. The adjusted census estimate for the census block is calculated by adding together the estimated adjustments for each post-strata represented in the block. Because of the problems of correcting a census with a survey, adjusted figures cannot be more accurate than the census counts in each of the 4,830,514 occupied blocks, or at all larger aggregations of them. There is no PES system—short of one which took a second perfect census—that could say adjusted counts are more accurate for all blocks. The question is whether the assumptions that underlie this synthetic adjustment mechanism are good enough to conclude that the counts are sufficiently accurate to be usable at a block or precinct level.

As noted above, the synthetic adjustment process rests on the assumption that persons in each post-stratum are homogeneous with respect to their probability of being missed by the census, *i.e.*, their capture probability. This is admittedly a very difficult thing to measure. There were several approaches taken by the Census Bureau to validate the homogeneity assumption, all contained in project P12.

The first part of P12 collapsed the 1392 post-strata by age and sex into 116 larger groups. To test whether the people living on blocks within these 116 larger post-strata are homogeneous with respect to capture probability, the Census Bureau conducted an analysis of the homogeneity of 115 of the 116 larger post-strata (the 116th is persons living on Indian reservations). A regression model predicted an adjustment factor for block parts, then compared that with an adjustment factor of 1.0 (no adjustment) representing the numeric census counts. This predicted adjustment factor was also compared with the measured factor for the post-strata used in creating the adjusted counts. For 24 of the 115 post-strata the census count was superior while for 91 post-strata the adjusted count was superior in terms of numeric accuracy. The Director interprets these findings to "give support to the accuracy of the selected PES adjustment model." [60] Regrettably, this evidence does not directly address the homogeneity issue. Like the uncorrected loss function studies this simply compares the census and the PES to yet a third estimate (the regression equation) whose quality or closeness to truth is unknown. This cannot be called a test or even a verification of the homogeneity assumption. To pursue this approach, allowance should have been made for the true variance in the regression estimates in a manner analogous to that done in the Undercount Steering Committee Addendum for the target population. It must be understood that such errors can easily occur when cutting edge research is used for production purposes under extreme time pressure.

The second part of P12 analyzed the homogeneity of state parts within post strata. Techniques known as analysis of variance were used to determine the validity of using post-strata, rather than states, for estimation of adjustment factors. The study was designed to determine if there was relatively more homogeneity within state or within post-strata. The study showed that, with the exception of the Mid-Atlantic Division, state differences were not significant within post-strata. This result was compatible with the conclusion that there is relative homogeneity for state parts within post-strata. There is no evidence of homogeneity for other geographic levels.[61] The only conclusion that can be drawn from this study is that the Census Bureau was better off using the actual post-strata for synthetic estimation than using any state-specific effects. Whether the levels of homogeneity within post-strata are acceptable is not even addressed.

The third part P12 looked at state homogeneity from a different vantage point. It measured whether other factors that are often correlated with undercounting are homogeneous within post-strata. Contrary to the results of the second part of the studies, these factors showed significant heterogeneity by state within post-stratum for well over 80% of the post-stratum groups. This study went further and measured the homogeneity of some of the components of the dual-system estimates at the block level and found about 14% of the post-strata groups to have significant state effects. Thus, the evidence in this study for the presence or absence of homogeneity within post-strata is mixed.

In summary, the analysis presented for decision from P12 was substantially different from that planned by the Census Bureau and used only the State as a surrogate for heterogeneity. We clearly do not thoroughly understand whether or not heterogeneity is a real problem. There are indications that using post-strata for synthetic adjustment is better than using states, but nothing more. It is impossible to conclude from any information the Bureau has presented in P12 that there is not residual heterogeneity within post-strata.[62]

Project P15 approaches the homogeneity problem by attempting to measure the quality of the dual system estimates by examining their expected variability. The measure used to do this is called the coefficient of variation which is the ratio of the sample standard deviation to the sample mean. The PES was designed so that these coefficients of variation were expected to be equal to 0.7 percent for the areas used in the design. In fact, in 48 of the 54 areas examined, the actual coefficients of variation are larger than expected. They ranged from 0.45 percent to 4.4 percent. This is direct evidence of substantially more variability in the DSE than expected and indirect evidence of heterogeneity within post-strata.

Other arguments have been made about this guideline. As noted in the analysis of Guideline One, Ericksen, Estrada, Tukey, and Wolter rely heavily on the Wolter/Causey/Tukey argument that synthetic adjustment will increase the accuracy of the counts.[63] For the reasons explained in the discussion of Guideline One, I do not find this argument compelling. Its reliance on the unsubstantiated homogeneity assumption simply emphasizes the concerns raised earlier.

Estrada argues that it is not necessary to show that the adjusted counts have to be better for all purposes, if it is shown on average to improve counts for its principal uses. "Improved counts to meet Constitutional needs for reapportionment and redistricting would be sufficient justification to adjust, even though for some other uses adjusted counts are less valid." [64] I do not consider this argument persuasive. Reapportionment and redistricting counts are the most demanding in terms of accuracy because block level counts are required to accomplish both.[65] If adjusted counts are better for these purposes, then they would necessarily be better for all others.

McGehee asserts that "variances between processing offices and evaluation strata fall outside expected levels in a number of the evaluation studies. At the district office level and below the data contain such wide variances that they could not be reconciled without weighting them to

---

yant, page 18.

[61] The Undercount Steering Committee report states that a majority of the Undercount Steering Committee believe this result would hold for other geographic levels. *However, there is no evidence presented to support this.* Undercount Steering Committee, page 2.

[62] Estrada agrees that the findings from P12 are mixed, although his conclusions differ from mine: "It supports the fact that PES poststrata are homogenous with respect to expected coverage error, but also questions the homogeneity of the division level poststratum. These findings lend support to the accuracy of the adjusted count based on the synthetic method particularly within poststrata and block-to-block variance in characteristics and net overcounts and undercounts. Overall, 79 percent of the time the adjusted count is better than the census count. Nonetheless, this research 'flags' the need to be aware of State effects." (Estrada, page 18.) Given the fact that reapportionment depends critically on state counts, Estrada's conclusions raise a large flag in terms of accuracy.

[63] See Estrada, page 29; Wolter pages 7–8; and Ericksen *et al.*, pages 20–21.

[64] Estrada, page 18.

[65] Recently, Mississippi's proposed redistricting plan was overturned by the Department of Justice for failure to use block-level data.

much higher levels." [86] As an example he notes that "the [matching] effectiveness rates varied from a low of 87.2% in Albany to a high of 93.49% in Kansas City. . . . [T]here was a significantly different level of success in Kansas City than in Albany. But why? The answer is that we do not really know." [87]

Special Advisory Panel Member Tarrance links the usability of adjusted counts for redistricting with the disruption the use of such counts would cause.[88] These arguments will be considered under Guideline Seven.

Wachter has serious concerns about the usability of these adjusted counts. I consider his concerns about state population totals and reapportionment under Guideline Three.[89] He does, however, present evidence that casts serious doubt on homogeneity within post-strata. Because "very little is known about local heterogeneity in census coverage," [90] he conducted simulations on 10 selected PES block clusters to determine the effect of an adjustment on both the improvement in the numeric level of the population at the district office level and the improvements in the shares or proportions of the population in a given district office. In other words, he considered both numeric and distributive accuracy. In Wachter's simulations, the level of the population is improved about twice as often as it is worsened by an adjustment. However, the shares suffer much more from the simulated adjustment. On average 59% of the office proportions are better, but the range over all the simulations shows anywhere between 39% to 78% improvements. Furthermore, in 7% of the simulation trials a majority of the districts are made worse. Now in any simulation, a true population for a block must also be simulated. Wachter argues that truth is chosen in his simulations so as to overestimate improvements achievable by an adjustment.

Wachter's evidence on heterogeneity is the only evidence that looks at actual behavior in the 1990 census and PES below the state level, and the only evidence that looks at the effect of heterogeneity on the shares of the population rather than the population levels. He states that his results are preliminary and need more work—but at least they are results that bear directly on the homogeneity issue. I find compelling his conclusion that "local heterogeneity is a serious problem for adjusting the 1990 census at district levels. My evidence indicates that a substantial portion, possibly a majority, of relative counts for district-size units can be made worse off by adjustment." [91]

Wachter made other efforts to measure block-to-block heterogeneity and district-to-district heterogeneity. These other attempts are inconclusive and neither support nor deny the homogeneity assumption, so, therefore, I did not consider them to weigh either for or against an adjustment.[92]

Heterogeneity and local variability pose a vexing problem for synthetic adjustment as GAO noted in their testimony.[93] In his article, Freedman makes this clear:

> Variability is a major obstacle to adjustment. Indeed, undercount rates differ from one geographical area to another, and from one demographic group to another. That is why synthetic estimates for small areas, based on demographic analysis, have not been widely accepted. However, adjustment by the DSE [Dual System Estimate] is unsatisfactory for the same reason. For example, one post-stratum consists of Hispanics—cross-classified by age, sex, and housing tenure—in central cities in the Pacific Division (California, Washington, Oregon, Alaska, and Hawaii). In round numbers, the 1990 population of the Pacific Division is about 40 million with 6 million Hispanics, 5 million of the latter being in southern California.
>
> Consider an adjustment for Stockton, a city of about 200,000 people in California's Central Valley, a 4-hour drive north of Los Angeles. The Hispanic population is about 50,000; there can be at most a few dozen Hispanics from Stockton in the PES [Post-enumeration survey], and a handful of gross omissions [persons counted in the "p" sample who were not in the "e" sample (census)] or erroneous enumerations [persons counted in the "e" sample (census) who were not found in the "p" sample]. No stable estimates could be developed from a sample that small. Instead, estimates for Stockton would be based on the adjustment factor for the whole post-stratum, the numbers being driven by PES data from southern California. The basic assumption: undercount rates for Hispanics are the same in Stockton as in Los Angeles. *There is no empirical evidence to support this assumption.* [Emphasis added.] And there is a similar problem for non-Hispanics. Indeed, adjustment factors for non-Hispanics in Stockton are driven by PES data on non-Hispanics in the whole Pacific Division. Apparently, Stockton's non-Hispanics are supposed to be like their counterparts in the north, while its Hispanics are taken to be southern. *Stockton is the rule not the exception.* [Emphasis added.] There are 39,000 state and local government areas to adjust: and only 5,000 sample blocks with PES data. *Most jurisdictions would be adjusted on the basis of data from elsewhere* [Emphasis added.]—and the synthetic assumption.[94]

None of the evidence I was given, other than Wachter's, confronted the measurement of this problem head on. The questions that remain unanswered are fundamental: What is the extent of residual heterogeneity within post-strata down to the county, city, precinct and block level, and what is the effect of that heterogeneity on the adjusted estimates both in levels and shares? Until this is known, the statement that the counts are usable for all census purposes is no more than an assertion.

Conclusions

I conclude that the considerations pointed to by Guideline Two tend to reject use of the adjusted figures and support use of the census counts. The adjusted figures—like the census counts—are consistent across all jurisdictional levels and of sufficient detail for all purposes. However, the adjusted figures do not appear to be of sufficient quality to be usable for reapportionment and redistricting. First, the distributive accuracy of the census counts is superior as concluded above in my review of the evidence on Guideline One. Furthermore, substantial evidence casts doubt on the homogeneity assumption underlying the entire synthetic adjustment methodology. Even if the tests discussed under Guideline One and based on the homogeneity assumption had proven favorable to adjustment, this evidence would weigh against adjustment. Instead, both considerations imply that the adjusted figures are not of sufficient quality to be usable for reapportionment and legislative redistricting. Thus, this Guideline weighs in favor of a decision not to adjust the census.

Guideline Three

The 1990 census may be adjusted if the estimates generated from the pre-specified procedures that will lead to an adjustment decision are shown to be more accurate than the census enumeration. In particular, these estimates must be shown to be robust to variations in reasonable alternatives to the production procedures, and to variations in the statistical models used to generate the adjusted figures.

---

[86] McGehee, page 32.

[87] McGehee, page 6.

[88] See appendix 2: Tarrance, V. Lance "Report to the Secretary of Commerce," Member, Special Advisory Panel, June 14, 1991, (hereafter Tarrance), pages 17–18.

[89] Wachter, pages 24–28.

[90] Wachter, page 26.

[91] Wachter, page 29.

[92] Wachter, pages 30–32.

[93] See the quotation from GAO in Guideline One.

[94] Freedman, pages 1233–1236.

Explanation

The Bureau of the Census will determine the technical and operational procedures necessary for an adjustment decision before the results of the post-enumeration survey are known. This procedure shall be chosen to yield the most accurate adjusted counts that pre-census knowledge and judgment can provide. The Bureau of the Census will then assess the components of systematic and random error in the procedure and it will assess the robustness of the estimates generated from that procedure. Various procedures and statistical models can be used to generate estimates of net overcounts or net undercounts and adjustment factors. This guideline specifies that a set of procedures for generating proposed adjusted counts will be determined in advance of receiving the 1990 post-enumeration-survey estimates. This guideline requires that these procedures be evaluated. These evaluations will identify other procedures and models that could be considered as reasonable alternatives to the chosen production process. These alternatives will be used to assess the accuracy and precision of the proposed adjusted counts. In addition they will be used to assess whether and by how much the adjusted counts could vary if alternative procedures were used.

Discussion

There are three questions raised by this guideline that have not already been dealt with in my conclusions about accuracy in the discussion of Guideline One:

(1) Were the procedures followed pre-specified?

(2) Were the estimates robust to production alternatives?

(3) Were the estimates robust to alternative statistical models?

Prespecification

The question of prespecification is difficult. No production of the complexity of the census or the PES can be completely prespecified. There are always unforeseen events that occur and that require modifications to the plan. In fact the procedures for the PES and for generating an adjusted count of the population were, broadly speaking, as prespecified. Even though there were several decisions, of some importance, made in the course of the estimation procedure, all were made solely by the career professional staff at the Census Bureau. The decisions reflected the best professional judgment of those career public officials vested with the responsibility for the census and the PES.

First, a decision was made not to combine DA with the PES to generate dual-system estimates. Second, there was a choice made of carrier variables to be used in the smoothing process. These variables help determine how the raw adjustment factors (published on April 18, 1991) are converted to the smoothed adjustment factors (published on June 13, 1991). Finally, in the smoothing process itself some observations which were either peculiar in their magnitude or their variance were treated specially. The Special Advisory Panel members were consulted in trying to deal with the difficulties encountered in the smoothing process.

Kruskal, Tarrance, and McGehee all raise concerns about the prespecification question. It is Kruskal's impression "that choice of the so-called smoothing procedures was profoundly based on PES results. One might indeed argue that such a choice has major merits, but it does not seem to me to follow the Guideline" [85] McGehee argues more strongly: "One's confidence is further eroded when—in an effort to explain unexpected results—the Bureau resorts to novel explanations, re-manipulation of the data, and a variety of other *ad hoc* techniques." [86] Tarrance expresses similar concerns: "Some procedures have been pre-specified but, as in all statistical operations, others have been suggested and/or adopted as the operations have been carried out. I have been concerned to note that a number of changes have been made in the last 18 months." [87] He also notes that "any attitude of 'if the numbers don't come out the way we think they should we can change plans' is diametrically opposed to what good government policy should allow. Furthermore, it is clear that the adjustment process is a statistical operation which has never been done before and there are many last-minute decisions being made." [88]

Ericksen, Estrada, and Tukey either find no problem with the prespecified procedures or do not mention it. Wolter notes that there were procedures in the enumeration that were changed late in the enumeration process that affected the PES; however, PES managers were able to cope with the changes in procedures. He also notes the decision not to combine the PES and DA and the smoothing decisions made during the PES process. He finds that each was treated with a high degree of professionalism.[89]

In any estimation process unforeseen difficulties will arise and no estimating system can be put on automatic pilot. The unsettling problem is that, as we will see below, the choices that occurred did make a difference in the outcome of the adjustment—differences large enough to change the implied apportionment of the Congress—and that different choices producing different results may have been made by other responsible individuals in the exercise of their best judgment. The enumeration process itself cannot be influenced in such a way. Any individual decision either has a tiny impact or is so distant from the final result (both in temporal terms and in statistical terms) that the decision maker does not know the import of his decision. This is simply not true of the types of decisions made here in the course of calculating PES count estimates. State counts were easily available to the persons deciding which smoothing method would be used. Although I believe that the decisions were made for sound professional reasons in the 1990 census, using these adjustment mechanisms opens the possibility for manipulation of future post enumeration surveys in ways that are unavailable in traditional census procedures. This weighs heavily against an adjustment of the census.

Robustness of the Results

I will combine the discussions of the robustness to alternative statistical methods and production methods in this section because they are for the most part intertwined.

One area in which statistical models could have an impact on the result of the PES is in the imputation of match status. As individuals from the PES are matched back to the census some cannot be definitively declared matched or unmatched, often due to missing data. The missing data were imputed to the unresolved cases and a match status was then assigned using a series of statistical models. The levels of missing data were sufficiently low that variation in these models made essentially no difference in the outcome of the PES (Studies P1 and P2). Here I concur with the Undercount Steering Committee judgment that the outcome is robust to the alternatives considered, although, as noted above, Wachter warns that unexamined assumptions underlie the statistical imputation models and, in

[85] Kruskal, page 4.

[86] McGehee, page 4.

[87] Tarrance, page 20.

[88] Tarrance, page 21.

[89] Wolter, page 9–10.

fact, the results could be sensitive to these assumptions.[100]

Wachter notes that the sensitivity of the imputation results to these unexamined assumptions, however, could have an impact on the apportionment of the House of Representatives that would be implied by an adjustment. He considers five alternative adjustment calculations: the smoothed estimates, the raw estimates, two of his imputation alternatives, and a fifth estimate that uses state adjustment factors based only on PES data gathered within that state. He finds that each method implies a different apportionment of the House, and eleven states either gain or lose a seat in at least one of the five alternatives. This instability in the results of the adjustment for the Constitutional purpose of the census argues strongly against an adjustment.[101]

The second area in which different methods could have affected the outcome is in poststratification. All the members of the same post-stratum receive the same adjustment factor. If post-strata are chosen differently then outcomes may be different. The Census Bureau investigated whether changes in the post-stratification by census division would change the results significantly by using an alternative post-stratification by state. This showed that three states would have had significantly different counts. It is also important to note that any variation due to uncertainty in post-stratification is not incorporated in the total error model.

A third area of concern is that of smoothing procedures. Smoothing is a technique that is used to remove some of the effects of random variability in the estimates of the adjustment factors for the 1392 post-strata, while preserving the meaningful systematic differences between subgroups. Since these adjustment factors are the results of a statistical process, they are subject to random variation. If you had taken a second *sample* the answer would be different. But some variation across the different poststrata is a result of real differences in behavior not simply random statistical variation. The point of the smoothing exercise is to remove the *random* variation while attempting to retain the real differences.

Smoothing involves three major judgmental decisions—the treatment of outliers, the variance pre-smoothing, and the choice of so-called carrier variables. We consider first the treatment of outliers. This is an extremely complex problem that posed great unforeseen difficulty for the Census Bureau. Let me start with a simple observation. When the final PES numbers were announced on June 13, 1991, a modified set of PES numbers was included as one of the alternatives considered as a possible set of final PES numbers but not selected. This set of numbers stood apart from the census and was closer to the selected method than the census. Thus it was a candidate for selection. This alternative, had it been chosen, would have implied a different apportionment of the Congress than the selected method. If the selected method were chosen and if the Congress were reapportioned on the basis of those numbers, California and Arizona would gain one more seat each and Pennsylvania and Wisconsin each would lose one seat compared to the census. Use of the modified PES estimates instead of the selected method would have resulted in a shift of only one seat—from Wisconsin to California. It is important to note that the only difference between the two methods is that, in the selected PES, 28 outlying variances out of 1392 variances were omitted from variance smoothing. In the modified version these 28 points were not omitted. Thus changing the treatment of only 2% of the points could have changed the allocation of one seat in the House of Representatives. I have included in Appendix 10 a list of State, county, and city populations under three smoothing schemes: the selected method, the modified method, and the raw adjustment without smoothing. Some of the sensitivities to smoothing choice are evident from these charts themselves. Let me highlight a few:

- The undercount rate for Arizona is estimated to be 2.8% under the modified PES smoothing scheme and 3.3% under the selected PES smoothing scheme.
- The undercount rate for Maryland is estimated to be 2.5% under the modified PES smoothing scheme and 1.8% under the selected PES smoothing scheme.[102]
- The undercount rate for the District of Columbia is estimated to be 5.8% under the modified PES and 5.0% under the selected PES smoothing scheme
- The undercount rate for Akron, Ohio, is estimated to be 2.2% under the modified smoothing scheme and 3.0% under the selected PES smoothing scheme.
- The undercount rate for Pasadena, Texas, is estimated to be 3.7% under the modified smoothing scheme and 3.0% under the selected PES smoothing scheme.
- The undercount rate for Miami, Florida, is estimated to be 5.4% under the modified smoothing scheme and 4.8% under the selected PES smoothing scheme.

The Census Bureau analysis emphasizes that the set of various population estimates derived from different smoothing methods are broadly similar in the counts they produce and, as a group, distinct from the census enumeration. I believe that, in fact, it would be difficult to choose on any objective statistical grounds among the host of alternatives the Census Bureau considered which do in fact produce different results for the Constitutional purposes of the census. As noted in the discussion of Guideline One, accuracy must be considered in terms of the distribution of the population not numeric accuracy. The Census Bureau analysis does not consider the similarity in terms of the population distribution of the sets of estimates or whether the variance inherent in those estimates, warrants the discarding of the census in favor of one of the particular estimates.

Wachter's analysis of the smoothing procedures that the Census Bureau used in developing the adjustment estimates also raises some serious concerns. He believes that "smoothing has turned out to be the most problematic part of the adjustment calculations," and that "the evidence leads me to fear that the smoothing has had more of an effect on the final adjustment than can be easily justified."[103]

As noted above, smoothing is a technique that is used to remove some of the effects of random variability in the estimates of the adjustment factors for the 1392 post-strata, while attempting to preserve the meaningful

---

[100] Wachter, pages 21–22.

[101] In connection with the loss function studies discussed in Guideline One, the Census Bureau compared the apportionment implied by the census to that implied by the so-called target population. They differed by two seats. The Bureau then considered 1000 random draws from the production DSE statistical distribution and compared the apportionment that would result from each draw to the target population apportionment. For 591 of the draws the production DSE apportionment did not differ from the target apportionment. For 367 of the draws there was a difference of one seat. For 42 there was a difference of two seats. This only shows that the PES estimator of apportionment differed from the target apportionment by 0.45 seats on average. It says nothing about the quality of the census, since the target is simply another adjusted estimate of the population, as the discussion of Guideline One demonstrates.

[102] If a state is estimated to have a greater than average (*i.e.*, 2.1%) undercount it gains proportionally from an adjustment. States below 2.1% lose. Thus the choice of adjustment method, had adjustment been used, would have determined whether Maryland was a winner or loser.

[103] Wachter, page 23 and page 24.

systematic differences. This is done [illegible] a technique called linear [illegible]ssion that "holds constant" [illegible]utes of the population we expect to be associated with low or high undercount rates in an attempt to isolate the random variation. The choice of the attributes to be "held constant"—also called carrier variables—is a matter of concern and will be discussed below. These regressions yield estimates of adjustment factors that supposedly have been purged of their random variability. Wachter characterizes these estimates as being "flattened." [104]

To calculate the smoothed factors one takes an average of the raw adjustment factor (before flattening) and the flattened adjustment factor—but a weighted, not a simple, average is used. For a particular post-strata, if you have observed a lot of random variability, the smoothed factor is chosen to be closer to the flattened factor—that is, the weight on the flattened factor is high and the weight on the raw factor is low. On the other hand, if the raw adjustment factor is fairly stable and does not show much random variation, you put more weight on the raw factor and less on the flattened factor when you calculate the smoothed factor. The smaller the random variation in a poststratum, the [illegible]re the smoothed factor relies on the [illegible]erved data and the less it relies on regression estimate.

But there is another level of complication. The measures of random variation, called variances, are themselves subject to random variation and, as happened in this PES, the variances can be large and unruly. The variances themselves vary a lot. When there are large measured variances, the smoothed factors are closer to the flattened estimates and on the whole, you tend to get lower adjustment estimates. The Census Bureau decided to soften this effect by pre-smoothing the variances before smoothing the adjustment factors. So there are two levels of smoothing—first variances, then factors.

Wachter shows that "the effect of deciding to use pre-smoothed rather than unsmoothed variances is to raise many of the adjustment factors by several percentage points and raise some by more than six percentage points. The changes introduced into the adjustment factors are of the same order of magnitude as the sizes of the adjustment factors themselves. These are huge changes for a decision of detail." [105] The fact that the statistical artifice of variance smoothing is making substantial differences in adjustment factors is disturbing. As Wachter observes:

> The raw adjustment factors are at only one remove from the data, the PES fieldwork that is the real information we have. Assumptions go into their computation and they are subject to many kinds of random and systematic errors. Notwithstanding these limitations, there is a fairly direct link between people missed or miscounted somewhere in a sample block and a big or small raw adjustment factor. Smoothing the factors themselves involves operating at two removes from the data, importing more assumptions, but incorporating information about variability that comes ultimately out of fieldwork. Pre-smoothing the variances that go into smoothing the adjustment factors is at three removes from the data. It incorporates little, if any, further empirical information. It depends entirely on another set of assumptions. [106]

The fact that pre-smoothing makes so much difference reflects the irregular and variable nature of the PES data. The implication is that the assumptions underlying the statistical models being used are important determinants of the outcome of the adjustment calculation.

Wachter discusses at length the reasons for variance pre-smoothing, but one argument he made was particularly striking. The variance pre-smoothing essentially results in large variances being made smaller and small variances being enlarged slightly. This seems to be the opposite of what is desired. A large variance means that the adjustment factor is not well estimated—it is noisy—so when smoothing the factor you should put *more* weight on the so-called flattened factor. Decreasing the large variance means you put *less* weight on the so-called flattened factor. The opposite argument can be made for small variances. Therefore, variance pre-smoothing is arguably having a result exactly opposite from that intended by the smoothing process. In addition, because low adjustment factors tend to have small variances, pre-smoothing makes those variances higher and thus systematically discounts the evidence of low net undercounts. [107] In other words, presmoothing tends to artificially inflate already high undercount rates and artificially dampen already low undercount rates. [108]

Wachter cites three other problems with variance pre-smoothing: First, the variance smoothing is not directed at making covariances more accurate. In his view the motivation for presmoothing was heuristic. Second, there are no strong reasons for choosing among the many models available to actually smooth the variances. Third, the choice about variance pre-smoothing affects not only the adjustment factors but the total net adjustments for broad aggregates of the population. For example, the variance preamoothing changes the estimated net undercount in the West South Central region from 2.95% to 2.76%. In the East South Central, it changes from 2.43% to 2.68%. [109] Again, these are changes of a very significant magnitude given the size of the national net undercount.

The choice of carrier variables in the statistical regression procedures used to smooth the adjustment factors could have a significant impact on the outcome. The Special Advisory Panel commissioned a study by David Hoaglin to study this impact. This study is used extensively in the arguments of Ericksen, Estrada, Tukey, and Wolter. The conclusion was much the same as with the various treatment of outliers. The carrier variable choice made a difference, although in absolute numeric terms not a huge difference. The 13 models Hoaglin produces look roughly similar to each other and to the production PES estimates all of which are distinct from the census. The same is true if relative shares for the thirteen evaluation post-strata are computed—the various carrier variables produce results closer to the production PES estimates than to the census. No results are available at finer geographic levels (such as states, counties, or cities.)

---

104 Wachter, page 35.

105 Wachter, page 38.

106 Wachter, page 37.

107 Wachter, page 30.

108 In their letter submitted on July 11, 1991, Ericksen and Tukey dispute Wachter's concerns over variance pre-smoothing and contend that variance pre-smoothing helped the accuracy of the adjustment. In his rebuttal letter, submitted on July 12, 1991, Wachter stands by his statements. Both letters are contained in Appendix 16. It is difficult to referee this dispute in the eleventh hour, especially since the lateness of the Ericksen/Tukey letter gave little chance for Wachter to prepare a detailed response. In checking with the Census Bureau, I have found that, in fact, the pre-smoothing operation was agreed upon in advance, but in mid-May difficulties were encountered in that operation. The Census Bureau consulted with the panel and Tukey offered several remedies that Hoaglin and Glickman refer to as "prescriptions in the spirit of a mustard plaster . . . not a tightly specified procedure derived from established statistical theory." (Appendix E of Ericksen, *et al.*, page 18.) Although Hoaglin and Glickman, seem to indicate that the choice among the three remedies should not have much effect on the ultimate smoothed estimates, one of the three shows exactly the phenomenon that concerns Wachter, of raising high variances and lowering small ones. (See Appendix E of Ericksen *et al.*, page 22.) In fact, the choice of the "mustard plaster" did have an effect on apportionment (see main text above). Finally as Wachter points out, there is disagreement as to what constitutes a reasonable alternative.

109 Wachter, page 39 and Table 3.1.

Wachter's assessment of the carrier variable selection is that "the effects of variable selection are not negligible but they are not a central issue." [110]

Ericksen points out that the total error model shows that the effects of the PES biases on population shares for the 13 large evaluation post-strata are small. In addition, he contends that his examination of the two estimates in the June 13, 1991, press release, shows the state population shares to be stable for the states that would gain or lose seats if the House of Representatives were reapportioned on the basis of adjusted counts. His reasoning is that the adjustments are larger than one or two times the standard error.[111] The difficulty with his reasoning is that it only considers sampling variability and ignores whether the shares are robust with respect to alternative statistical and production methods.

Conclusions

I have previously concluded that the adjusted figures have not been shown to be more accurate than the census enumeration. That is all that is required under Guideline Three to conclude that the census may not be adjusted. There are, however, additional considerations under Guideline Three under which I also conclude the 1990 census should not be adjusted.

It has proved virtually an impossible task to prespecify the adjustment procedure. It is equally impossible to prespecify the Census procedure. However, in the adjustment procedure an individual or responsible group must make choices which have politically significant effects on the counts that can be transparent to those making the choices. This puts the counts at greater risk of being manipulated than the census. There is no evidence of unprofessional or political manipulation in the 1990 PES program.

The results of the adjustment procedure are broadly robust at an aggregate, national level. However, although various alternatives seem to distribute counts in roughly similar ways, small changes in methodology can move seats in the House. It is also true that small changes in the census enumeration can move seats in the House as well, but no individual involved in the enumeration process can predict how. That is not true for the decisions for adjustment that cannot be or were not prespecified.

One of the most problematic parts of the adjusted process was the bundle of statistical techniques contained in the smoothing process. These techniques relied heavily on statistical assumptions, resulted in large changes in adjustment factors, and may very well have led to an overstatement of the undercount. Thus, this guideline weighs in favor of a decision not to adjust.

[110] Wachter, page 41.

[111] Ericksen, page 3

Guideline Four

The decision whether or not to adjust the 1990 census should take into account the effects such a decision might have on future census efforts.

Explanation

The Decennial Census is an integral part of our democratic process. Participation in the census must be encouraged. Respect for the objectivity, accuracy, and confidentiality of the census process must be maintained. Accordingly, if evidence suggests that adjustment would erode public confidence in the census or call into question the necessity of the population participating in future censuses, then that would weigh against adjustment.

On the other hand, if evidence suggests that the failure to adjust would erode public confidence in the census and thus result in widespread disinclination to participate in future censuses, that would argue for adjustment. The extent to which a non-adjustment would be perceived as a politically motivated act, and thus would undermine the integrity of the census, should also be weighed in making any adjustment decision.

Discussion

There is no scientific or quantitative means by which we can determine with reasonable certainty the impact of a decision made in 1991 on human behavior and activities that will occur in the year 2000 and beyond. Indeed, this guideline merely requires that we consider the effects that our decision today might have on future census efforts. In my view, such consideration requires that we examine relevant information and draw upon past Census Bureau experience as well as common sense in making rational predictions about such effects.

The universe of "future census efforts" encompasses a wide variety of activities: the efforts of individuals in completing census forms and cooperating with enumerators; the efforts of state and local officials, civic leaders and special interest groups in supporting outreach programs, public awareness campaigns, and active involvement in counting their target populations; the efforts of Census Bureau workers in enumerating as many households as possible; the efforts of Census Bureau professionals in making judgments and decisions about procedures to achieve the most accurate counts possible and to ensure objectivity and integrity of the process; and the efforts of the Department of Commerce, which includes the Census Bureau, to ensure appropriate levels of funding from Congress to support its enumeration activities. Each of these activities affects participation in and coverage of the census. To the extent that we can draw on relevant data, observations, and experience, consideration of the effects of decisions to adjust or not adjust on each of these activities is appropriate.

Sources relevant to our considerations include a study by the National Opinion Research Council (NORC), [112] public comments on the adjustment decision,[113] comments on Guideline Four submitted by the members of the Special Advisory Panel,[114] and discussions with experienced Census Bureau officials. Based on these sources, it is my conclusion that there is greater risk of potential harm to future census efforts as a result of a decision to adjust than as a result of a decision not to adjust. A discussion of the possible effects on each of these activities follows.

Effects on Individual Participation

Recently, the Census Bureau commissioned a study by the National Opinion Research Corporation (NORC) to try to measure how an adjustment might affect future census behavior by means of a telephone survey of a representative national sample of 2,478 households.

Persons were asked to evaluate the likelihood that they would participate in the next census. Then they were asked how that likelihood would change if there were an adjustment and how that likelihood would change if there were not an adjustment. The results were paradoxical—both a decision to adjust and a decision not to adjust would decrease the likelihood of participation.

The survey shows that the adjustment issue is not high in public consciousness or well understood. Only one-quarter (23.4 percent) of persons said they had seen or heard anything about the census in the past few months. When probed about what they had seen or heard, only 14.1 percent spontaneously mentioned anything to do with adjustment, undercount or errors in the census count. When told that people are talking about whether or not to adjust the

[112] See Appendix 11. National Opinion Research Corporation. The Potential Impact of Adjusting of Not Adjusting the 1990 Census. June 19, 1991.

[113] See summary of comments on Guideline Four in Appendix 8.

[114] Ericksen, page 2; Estrada, page 20; Kruskal page 4; McGehee, page 33; Tarrance page 23; Tukey, page 2; and Wachter, page 42.

results of the census to correct for errors in counting the population, 22.3 percent recalled they had seen or heard something about this. Probing questions showed that only 4.9 percent understand the adjustment issue.

Prior to any discussion of adjustment, a total of 84 percent of those surveyed stated they were "extremely or very likely to participate" in the next census. After the discussion of adjustment, 75.5% were "extremely or very likely to participate" in the future if the census were adjusted, as compared to 71.3% if it were not. Thus, while these results indicate that intention to participate in future censuses is marginally higher if the census were adjusted than if it were not, there is less inclination to participate in the future regardless of the outcome of the decision. As NORC points out in its conclusions: "While large numbers remain very favorably disposed to participating in the next and future censuses, this intention is a very slippery, ephemeral and changeable one . . . subject to influence by factors like the adjustment decision or, more likely, from the controversy or fallout emanating from the events that follow that decision." The survey also indicates that, prior to any discussion of adjustment, 5.5 percent were "not very likely" to participate in the next census. A decision to adjust would result in 5.3 percent in the "not very likely" category. A decision not to adjust would result in 8.6 percent in this category.

It is unclear what this survey meaningfully demonstrates, other than confusion over what an adjustment is and the negative effect of the controversy over adjustment on the present perception of a person's likelihood of participation in future censuses. However, as the survey report emphasizes, the need to explain the issue of adjustment and its implications will necessarily outlive the survey and the adjustment decision itself, and the inability of the surveyors to explain the issues to those surveyed is certainly grounds for some caution.

The division of public opinion on the future effects of adjustment indicated by this survey is consistent with the division of opinion demonstrated by the public comments received by the Department. While some claimed that an adjustment would erode public confidence in the census and thus lower future participation, others claimed that a decision not to adjust would erode public confidence and thus lower future participation.

The explanation of this Guideline states that evidence of widespread disinclination to participate in future censuses as a result of a decision not to adjust would weigh in favor of an adjustment. Neither the public comments nor the NORC survey provide evidence that this will occur. Indeed, the NORC study indicates that a decision not to adjust would make only 8.6 percent "not very likely" to participate in the future, just 3.1 percent more than those who would be "not very likely" to participate in any event. Thus, while there would be some additional disinclination to participate, it would not be widespread.

The explanation of this Guideline also states that evidence that calls into question the necessity of the population participating in future censuses would weigh against an adjustment. A number of the public comments express concern that an adjustment would result in the perception that an individual's failure to participate would be compensated by an adjustment and thus lower participation. In light of this, I am skeptical rather than optimistic about the likely motivation of individuals to participate in the future if an adjustment were made. However, I do not find compelling evidence in either direction regarding the effects of a decision on future individual motivations.

Effects on Complete Count Efforts by State, Local, Civic, and Interest Group Leaders

A number of the public commentators, as well as Wachter [115] and Tarrance,[116] expressed serious concerns that an adjustment would negatively affect the efforts by state, community, civic, and interest group leaders who traditionally provide essential support in encouraging participation in the census. I share these concerns. Currently, it is in the interests of every governor, mayor, and interest group to help get their target populations counted. The Census Bureau works closely with such officials and groups for two to three years before census day. The efforts include mapping, address compilation, massive advertising campaigns, and public awareness activities. I agree with my advisors who believe that such cooperative efforts are absolutely critical to the Census Bureau's mission to conduct an actual enumeration of all persons residing in the United States on census day, and particularly critical in reaching the hardest to count. Like others, I am concerned that an adjustment will remove the incentive that these public officials and groups currently have to provide active support in achieving a complete count.

[115] Wachter, page 42.

[116] Tarrance, page 23.

Based on the public comments, it is clear that many public officials believe that an adjustment will correct specific errors they have identified in the count of their communities. With such mistaken impressions, it is unrealistic to expect these leaders to put census outreach efforts above the many other claims on their limited resources. As Wachter predicts, complete count committees, local advertising, celebrity appearances, and special programs to ensure more complete minority counts would be likely to suffer as a result of an adjustment.[117]

Senior officials at the Bureau, including the Director, agree with this assessment. At the same time, the Director believes that states and cities will still have an incentive to encourage participation in order to get the best possible city planning data. I find this unpersuasive in light of the numerous public comments received from local officials demonstrating a profound lack of understanding of the effects of an adjustment and a misplaced faith in its ability to correct particular problems they have identified in their communities.

I find no evidence indicating that local support would decrease as a result of a decision not to adjust the census.

Effects on Funding of Future Censuses

Tarrance [118] and Wachter [119] expressed concern that an adjustment would adversely affect the Department's ability to obtain sufficient funding for future censuses. I share this concern. The most expensive element of the census is the extraordinary effort to count the last five percent. With the illusory prospect of an adjustment to achieve a full count in congressional districts and states, it would simply be unrealistic to expect Congress to appropriate funds to the full extent necessary to complete an enumeration of the hard to count. Without the funds needed to complete an enumeration, the quality of census data, especially in smaller areas, would be jeopardized. There appears to be little risk that Congress would deny such funds as a result of a decision not to adjust.

Effects on Efforts by Census Enumerators

As Wachter recognized, the future effects of a decision to adjust could be most severe on those temporary workers who must actually conduct the enumeration process.[120] The difficulties

[117] Wachter, page 42.

[118] Tarrance, page 23.

[119] Wachter, pages 42–43.

[120] Wachter, page 43.

of hiring, training, and supervising the thousands of temporary census employees are well-known and well-documented. It is time-consuming, often tedious, and occasionally dangerous work that requires extraordinary diligence for less than commensurate pay. There is a real risk that, with an expectation of a correction through adjustment, the field staff would not have the same sense of commitment and public mission in future censuses and, as a result, careless and incomplete work would increase, thereby decreasing the quality of census data. These are the workers the Bureau depends on to collect the data from the groups that are hardest to enumerate. If these data suffer, the information lost at the margin is information that is especially important to policy development.

I am unaware of any concerns that census enumerators would be less motivated as a result of a decision not to adjust the census.

Effects on the Independence of Bureau Professionals and the Integrity of the Census

Senior Bureau officials as well as Tarrance [121], Wachter [122], and McGehee [123] have raised concerns about the potential for manipulation of an adjustment for partisan purposes. As Wachter recognized, adjustment may pose significant risk to the technical independence of Census Bureau professionals who have traditionally been free from external influence in the implementation of their mission.[124] A principal drawback of adjustment is the fact that a few technical decisions can swing the outcomes of apportionment, redistricting, and Federal funding allocation. Decisions that may be nearly equally defensible from a technical standpoint may have very different outcomes which can be known in advance of the decisions. Thus, adjustment opens the door to manipulation of the census for partisan gain. It would therefore greatly increase not only external scrutiny and second-guessing of Census Bureau professionals and prospective candidates for key technical positions, but also inevitably increase pressure to politicize these positions. This would impose an even greater burden on technical staff in their attempts to make scrupulously objective and fair decisions. These risks pose serious threats to the integrity and objectivity of future censuses.

[121] Tarrance, page 5.

[122] Wachter, page 44.

[123] McGehee, page 33

[124] Wachter, page 44.

Concerns have also been expressed in the public comments and by Wolter [125] that a decision not to adjust the census may be seen as politically motivated and therefore adversely affect the integrity of the census. While I recognize these concerns, I believe they are outweighed by the likely adverse effects on future census efforts from an adjustment.

Conclusion

Based on the information available, I conclude that an adjustment would adversely affect future census efforts to a greater extent than any adverse effects of a decision not to adjust. The evidence indicates that the controversy over adjustment is likely to have a negative effect on future censuses regardless of the outcome of the adjustment decision. I am concerned that an adjustment would reduce state and local support for future censuses, adversely affect the Department's ability to obtain appropriate funding for future censuses, adversely affect the quality of the work done in the future by temporary census enumerators who are essential in reaching the hard-to-count, subject the Census Bureau to partisan pressures, and create the possibility for political manipulation of future census counts. Thus, this guideline weighs in favor of a decision not to adjust.

Guideline Five

Any adjustment of the 1990 Census may not violate the United States Constitution or Federal statutes.

If an adjustment would violate Article I, Section 2, Clause 3 of the U.S. Constitution, as amended by Amendment 14, section 2, or 13 U.S.C. section 195, or any other constitutional provision, statute or later enacted legislation, it cannot be carried out.

Discussion

In addition to the technical and operational aspects of the census and the proposed adjustment which I have considered in connection with Guidelines One through Four, I have also considered the constitutional and statutory implications of an adjustment decision. In my view, neither the Constitution nor the relevant statutory provisions are themselves conclusive as to whether the proposed adjustment would be unconstitutional or unlawful because the *sine qua nons* of constitutionality and lawfulness and the propriety of adjustment are the same: the need for unambiguous accuracy of the adjustment methodology and data. Because analysis of the significant legal issues is thus dependent upon the statistical analysis, which itself mandated my decision on the substantive merits not to adjust, it was unnecessary to decide the legal issues. This Guideline therefore only served to verify, not determine, my decision.

[125] Wolter, page 11.

Constitutional Considerations

While not free from doubt, it appears that the Constitution might permit a statistical adjustment, but only if it would assure an accurate population count. See *Carey* v. *Klutznick*, 508 F. Supp. 404 (S.D.N.Y. 1980); *Young* v. *Klutznick*, 497 F. Supp. 1318 (E.D. MI 1980). By implication, then, a determination that the proposed adjustment would not discernably or reliably improve the accuracy of the headcount would raise uncompromisable constitutional concerns, inasmuch as adjustment would not be contributing to the most accurate count, but rather would be injecting additional uncertainty and error. Thus, while the Constitution might not, *per se*, bar an adjustment, the question of whether a particular adjustment is constitutionally valid can only be made after the final form of the adjustment is known.

This principle—that an adjustment must be consistent with the constitutional requirement of "enumeration," i.e., an accurate count free from politicization and equivocation—is also supported by the intent of the Framers that the census utilize verifiable methods which obtain an accurate population count. This goal of accuracy would not be met, to give the clearest example, by mere guesswork. The central question under the Constitution thus supports, though it did not determine, my conclusion: the need for verifiable methodology and unambiguous data are the modern-day requisites of what was explicitly desired by the Framers when they provided for an "actual Enumeration." That phrase commands for all time that what shaped the details of the very first congressional apportionment (there was then as yet no census)—guesswork and political deal-making—never would be permitted again.

As the discussion of Guideline One demonstrates, the evidence of improved accuracy resulting from the proffered adjustment methodology is at best mixed. That evidence is not sufficient as a matter of substantive merit and, derivatively, it also fails the test prescribed under the Constitution. While the essence of my decision not to adjust rests in the uncertainty of the proposed adjustment and the questionable nature of the data produced, that very uncertainty and question mark the rough ...ls of politicization that the framers ...ated be avoided when they ...ed "enumeration," that is, an objectively accurate count.

Census Act Provisions

The Census Act contains two provisions authorizing the Secretary of Commerce to use sampling to conduct the decennial census. See 13 U.S.C. section 141(a) and 13 U.S.C. section 195.

Section 141(a) provides in pertinent part:

The Secretary shall, in the year 1980 and every 10 years thereafter, take a decennial census of population as of the first day of April of such year, which date shall be known as the "decennial census date", in such form and content as he may determine, *including the use of sampling procedures and special surveys.* (Emphasis added.)

Section 195 provides:

Except for the determination of population for purposes of apportionment of Representatives in Congress among the several States, the Secretary shall, if he considers it feasible, authorize the use of the statistical method known as "sampling" in carrying out the provisions of this title (Emphasis added.)

While judicial opinion is unsettled on the question of whether adjustment violates section 195, the majority of ...ts considering this issue have ruled ...section 195 permits an adjustment if ...adjustment method makes the census more accurate. See *Cuomo* v. *Baldrige*, 674 F. Supp. 1089 (S.D.N.Y. 1980), *Carey* v. *Klutznick*, 508 F. Supp. 404, at 415 (S.D.N.Y. 1980); see also, *City of Philadelphia* v. *Klutznick* 503 F. Supp. 663 at 679 (E.D. PA 1980); *City of New York et al.* v. *United States Department of Commerce, et al.*, (S.D.N.Y. 1990). *But see Orr, et al.* v. *Baldrige, et al.*, U.S.D.C., S.D. Ind., No. IP 81–604–C, July 1, 1985. Even assuming that the statute does not *per se* prohibit an adjustment, not all forms of adjustment would be sanctioned and the legality of the adjustment could only be determined after the form of adjustment is chosen. Thus, as with the constitutional issues, the analysis of the statutory issues cannot be separated from the analysis of the accuracy of the chosen adjustment method. Because the evidence of improved accuracy from an adjustment is insufficient, the standard articulated by the majority of these courts is not met. While this legal conclusion was not dispositive, it affirms my decision not to adjust based on the substantive merits.

Conclusion

...he question whether the chosen ...hod of adjustment would violate the Constitution and federal statutes depends upon the substantive analysis of whether accuracy of the census is improved by an adjustment. Because there are other compelling substantive reasons not to adjust, legal considerations did not provide a basis for my decision.

Guideline Six

There will be a determination whether to adjust the 1990 census when sufficient data are available, and when analysis of the data is complete enough to make such a determination. If sufficient data and analysis of the data are not available in time to publish adjusted counts by July 15, 1991, a determination will be made not to adjust the 1990 census.

Explanation

It is inappropriate to decide to adjust without sufficient data and analysis. The Bureau will make every effort to ensure that such data are available and that their analysis is complete in time for the Secretary to decide to adjust and to publish adjusted data at the earliest practicable date and, in all events, not later than July 15, 1991, as agreed to in the stipulation. Note, however, that the Department and the Bureau have consistently stated that this is the earliest possible date by which there is a 50 percent chance that an analysis could be completed on which a decision to adjust could be based. If, however, sufficient data and analysis of the data are not available in time, a determination will be made not to adjust the 1990 Census. The coverage evaluation research program will continue until all technical operations and evaluation studies are completed. Any decisions whether to adjust other data series will be made after completion of those operations.

Discussion

In order to evaluate the quality of the census and the post-enumeration survey, the Census Bureau conducted an extensive and ambitious research program designed to provide timely information on which to base a decision by July 15, 1991. Due in part to some unexpected anomalies in the data, progress on the evaluation was delayed in the final critical weeks, leaving the Bureau little time to complete its analyses. These pressures may have affected the quality of the research, and there is still much that we do not know about the quality of the PES and the adjusted counts relative to the enumeration. Nonetheless, based on the record available, I believe there is sufficient evidence to make a decision on adjustment.

The Census Bureau has done a remarkable job of condensing into a few short months a challenging evaluation program that was comparable to a multi-year research program for the 1980 census and the 1987 test of adjustment-related operations. The Census Bureau produced highly technical research on a very tight production schedule, using tools that were on the cutting edge of statistical theory and survey methods. The dedication, professionalism and hard work of Census Bureau staff under often intense pressure is truly commendable.

Although sufficient data are available for me to decide the adjustment question, it is important to note that because of the court imposed deadline for the decision, the analyses of the data are far from compete. All parties involved in the decision making process have expressed a desire for more time to digest and analyze the voluminous material created by the research program. I am particularly concerned about problems in data quality and analysis that were revealed, or occurred, in the final weeks before the decision.

Good research requires a careful weighing of the evidence, especially when it is on the frontier of the science. When such novel research is to be used for such far-reaching policy purposes, it requires discussion with peers who have not been intimately involved with the details so that some perspective can be gained. It benefits from probing questions, from looking at the data from different perspectives, from the use of alternative models and from intense and independent professional scrutiny. The time schedule simply did not permit a full range of such activities.[188]

Before the release of the selected and modified PES numbers, an inconsistency was found in the calculation of the margins of error upon reviewing the proposed release in its penultimate form. This was not a subtle error, but one that should have been caught by a careful cross-checking of the tables. After being informed of the inconsistency, the Census Bureau began work to discover its source. Fortunately, no fundamental error had been made. However, the release was delayed by almost two weeks, setting back an already tight schedule in the last critical weeks of evaluation. Such errors were the result of too much work being compressed into too little time. To its credit the Census Bureau worked hard

188 Kruskal makes a similar point (Kruskal, page 8) as does Tarrance (Tarrance, page 27).

to find the error, fix it, and ensure that accurate data were released.

Later, in reviewing the work of the Undercount Steering Committee, fundamental questions were raised about measurement of the relative accuracy of the census and the PES. The loss function analysis was found to be unconvincing. The Census Bureau was therefore asked to revisit parts of its work. As a result of these questions, the Bureau staff found an error in the calculation of the loss functions. Correction of this error changed the number of States for which the census counts were judged more accurate than the adjusted figures from 11 to 21—a substantial and significant increase.[137]

An Addendum to the Undercount Steering Committee report was filed on Thursday, June 27, 1991. In section 4 of that addendum, which is included as appendix 5, the Undercount Steering Committee states the following:

> Given this new information, the Undercount Steering Committee members reevaluated their positions regarding the report issued on June 21, 1991 * * *
>
> The new information added uncertainty to the decisions of the majority, but their overall conclusions were not changed. In addition, particular sections of the report present representations of committee opinions that are now weakened by the new information. The sections of the report most affected by these new data are:
>
> The statement on page 6 of the report that 29 of 50 States are made more accurate by adjustment would be changed under the new loss function analysis; and
>
> Page 4 of the report summarizes the conclusions of the committee regarding Guideline One. The summary indicates that the majority of the committee relied on the loss function analysis that showed that a large majority of areas were made more accurate by adjustment. This is a stronger statement than the position now held by many of the committee members.
>
> In conclusion the overall committee position has not changed regarding adjustment, but has been weakened somewhat. These new data also underscore the points raised in report's findings on guideline 6 (see p. 12-13). When additional information, as the data presented above, becomes available, the committee acknowledges that it may strengthen or weaken its conclusions. On June 21 the committee judged that further analysis would be unlikely to change its conclusion. The majority stands by its original conclusion while acknowledging that the ongoing work, had it been available by the date our recommendation was due, may have caused different "weighing" of the results.[138]

These eleventh-hour findings weakened a key piece of evidence favoring adjustment. Because of these two significant errors, my concerns about the sufficiency of data and the strength of analysis supporting adjustment were heightened.

A second example of the pressures of the schedule is that as of the afternoon of Thursday, July 11, 1991, just two working days before my decision would have to be announced, I received a communication from Ericksen and Tukey taking issue with some of the conclusions in Wachter's report. Although I understand that many of the issues surrounding adjustment will be debated for a long time to come, the fact that some of the members of the Special Advisory Panel feel it incumbent upon themselves to offer last minute advice reinforces my perception that a full professional airing of issues has not taken place. Wachter wrote a speedy response to Ericksen and Tukey which I received on Friday, July 12, 1991. But a last minute debate by letter is not the way to carry out the important dialogue required on these issues.[139]

Over the course of the next months and years the data will be studied, the models tested, the professional discussions joined. We do not know what will be discovered about the quality of the PES data and the models that led to the adjusted counts. I am sure that the Census Bureau will not compromise its richly-deserved reputation for thorough and careful research. We need those efforts to build toward a better census in the year 2000.

But the question is whether we should adjust the census based on the data and incomplete analysis that we have now. As Wachter notes, we must "strike a sensible balance between the need to reach closure and the need to check and study further."[140] The decision must be made on its merits.

Notwithstanding my concerns about the effect the July 15, 1991, deadline had on research efforts, I conclude that sufficient data exist to permit me to decide whether to adjust the census. I conclude that the data support a decision not to adjust. Among the facts that weigh against an adjustment are:

- The PES missed a significant number of persons whom we cannot locate. Thus we cannot judge whether the adjusted census is distributionally superior to the enumeration simply by putting back into the count those we can locate by the PES.
- At the most aggregate level, the PES would move the count of the population in the opposite direction for some demographic groups as compared to those implied by DA.
- There is no convincing evidence to suggest that the adjusted counts give a more accurate count of the distribution of the population across various levels of geography. In fact the evidence indicates the census counts probably yield a more accurate measure of the distribution of the population.
- There is no convincing evidence that homogeneity within the poststrata used in adjusting the census counts is a statistically valid assumption.
- There is evidence that the estimates of the population produced by adjusting the counts are sensitive to small changes in the estimation procedure and these have significant effects.

Thus I find that the evidence presented is sufficient to conclude that the counts should not be adjusted.

Conclusion

An adjustment to the census is a fundamental change in the way we count and locate the persons residing in the United States. I am deeply concerned that if an adjustment is made, it would be made on the basis of research conclusions that may very well be reversed in the next several months. That would be bad for the country and bad for the Census Bureau.

The results of the PES evaluation studies are not yet completely analyzed. Because of the compressed time schedule imposed by the July 15 deadline, the analysis has not been subject to the full professional scrutiny that such important research requires and deserves. To the Census Bureau's great credit, the statistical tools used to calculate and evaluate the adjusted counts are at the cutting edge of statistical research. But such cutting edge research is not tried and true—it requires more thorough scrutiny before it can be used to affect the allocation of political representation and Federal funding.

Nonetheless, the demands of good research must be weighed against the need for a timely decision. In time we may find a way of combining the PES and the census to create counts that better reflect the absolute levels and the distribution of the population. There are sufficient data and analysis to support a decision not to adjust.

Guideline Seven

The decision whether or not to adjust the 1990 Census shall take into account the potential disruption of the process of the orderly transfer of political

---

[137] As noted in Guideline One, these numbers are for the version of the analysis in which it is assumed that the measured variance is the whole story. As discussed there, the change is even more dramatic (from 11 to 28) if the true variance is assumed to be twice the measured variance.

[138] Addendum, page 6-7.

[139] Both letters are contained in Appendix 26.

[140] Wachter, page 46

representation likely to be caused by either course of action.

Explanation

This guideline is intended to ensure that the factor of disruption of the process of the orderly transfer of political representation is explicitly taken into account as the decision is reached. For example, many states have pointed to adjustment as being disruptive to their redistricting plans. Likewise, members of some communities that are believed to have been historically undercounted contend that if the Census were not adjusted, this would disrupt the orderly and proper transfer of political representation to their communities. The inability to ensure accuracy of counts at local levels may result in politically disruptive challenges by localities to official census counts.

This guideline recognizes that the Decennial Census plays a pivotal role in the orderly redistribution of political representation in our democratic republic. The process used to generate the required counts must not be arbitrary either in fact or appearance. The Secretary is thus obliged to consider the impact of his decision on the fairness and reasonableness of that redistribution to all those affected. This guideline requires an explicit statement of how and to what degree adjustment or non-adjustment would be disruptive. Even though these are concepts that are not easily quantifiable, they warrant serious consideration in order for the Secretary to make a prudent decision on an issue that profoundly affects public policy.

Discussion

Among the primary purposes of the census are to provide the basis for the reapportionment of the House of Representatives and the drawing of new Congressional district lines within states. Census figures are also used by most states to redraw state legislative district boundaries, as well as by cities and counties in redrawing their own districts.

The Clerk of the U.S. House of Representatives has officially certified to each of the fifty states the number of seats allotted to the state for the 103rd Congress based on the census figures released on December 26, 1990. As of May 1991, some 20 States had already enacted either or both of their Congressional and State legislative redistricting plans. The U.S. Department of Justice is reviewing approximately one dozen of the state plans as well as those of many cities and counties to ensure compliance with the requirements of the Voting Rights Act.[181]

If adjusted census counts were issued, Congress would have to decide whether to change the apportionment for the 103rd Congress which is to be elected in November 1992. If there were a decision to change the apportionment using the formula in current law, the Clerk would have to issue new certificates to the states advising them of the number of seats to which they would be entitled based on adjusted counts. If this change were made, the States of California and Arizona would gain one seat each and the States of Wisconsin and Pennsylvania would each lose one seat relative to the apportionment previously certified by the Clerk of the House.

It is unclear whether Congress would change the apportionment even if adjusted counts were chosen. The requirements of the statutes governing apportionment were fully met in January with the certification of the number of seats to each state. Thus, as noted in a number of public comments [182], additional action may be required on the part of Congress to change that apportionment. Whether, how, and when that action would be taken is for the Congress to determine.

It is important to remember, however, that the modern apportionment process was designed to be automatic. Once the counts were transmitted by the President to the Congress, the apportionment took place without legislative action. This design was intended to put an end to the blistering fights over apportionment that occupied earlier Congresses and, in fact, prevented reapportionment after the 1920 census, depriving citizens of a fair allocation of political representation throughout the nation for the remainder of the decade.[183] The adjustment of the Census might well create similar bitter disputes and paralyzing legal challenges over the apportionment of the 103rd Congress. The political implications of this are matters of substantial concern.

If the adjusted census were the basis for a reapportionment of the House, for the first time, the apportionment would not be determined solely on the basis of the number of persons within each State's border. This is due to the effects of cross-state groupings of post-strata in the PES on the adjustment process. For example, if the counts were adjusted, the certified population count for Delaware would depend on the results of the PES in Maryland, the District of Columbia, West Virginia, Virginia, North Carolina, South Carolina, Georgia and Florida. This is because Delaware is in the South Atlantic census division, and PES estimates are developed division-wide.

At the State level there is also likely to be confusion, disruption and extended litigation if the census figures are adjusted. Members of the Special Advisory Panel reported on extensive testimony they received from members of the National Conference of State Legislators in Baltimore, Maryland on June 28, 1990.[184] The testimony focused on the effects of an adjustment on the states' ability to accomplish redistricting in compliance with state-imposed legal deadlines. Witnesses were concerned that the electoral process would be paralyzed by the endless litigation which two sets of census numbers would be certain to provoke. Witnesses cited major problems with adjustment: costs and delays in drawing new plans, costs of additional elections, the need for costly special legislative sessions, time constraints, and charges of partisan tampering with census data. Based on the testimony, it is clear that adjustment would create serious disruption for at least a dozen states that have early redistricting schedules or constitutional deadlines. Some states have simply delayed starting the process until after the adjustment decision. As Estrada recognized, adjustment also would require modification of recently designed districts to meet one-person, one-vote requirements.[185]

Protracted legal battles that preclude redistricting in time for the 1992 elections would deprive minority groups and others the opportunity to realize and benefit from the gains achieved through demographic shifts during the past decade. The same pattern would likely occur in redistricting efforts for city and county elections, which have already begun in a number of areas. Moreover, the adverse effects of an adjustment on the accuracy of small area counts (as demonstrated in the discussions of Guidelines One through Three) would likely result in politically disruptive challenges by localities to adjusted counts.

Several public commentators, as well as Tukey,[186] noted that such disruption

---

[181] See appendix 12. Turner, Marshall, "Planning the 1990 Census Redistricting Data Programs," U.S. Bureau of the Census, [hereafter Turner].

[182] See the summary of public comments on Guideline Seven in appendix 8.

[183] See the discussion of this matter in Chapter Six of Margo J. Anderson, The American Census: A Social History. New Haven and London: Yale University Press. 1988.

[184] Tarrance, page 28 and Wachter, page 47.

[185] Estrada, page 23.

[186] Tukey, page 2.

was foreseeable at the time of the Department's decision to consider an adjustment and that anticipated effects should not be considered in making the decision. The fact that disruption could be anticipated does not mean that it should be ignored. Indeed, consideration of disruption as a factor to be weighed in the decision was legally upheld. Moreover, as Tarrance stated, "we would not be responsible stewards of the public trust if we do not understand that we are considering more than just a scientific statistical improvement of an imperfect government program." [137] Because the census is the basis for allocating political representation in our country, the public policy implications of adjustment, including resulting political disruption, had to be considered in reaching this decision.

The potential for disruption as a result of an adjustment must be weighed against any disruption that would occur from a decision not to adjust. There will inevitably be litigation resulting from a decision not to adjust that may also delay or disrupt redistricting. Some public commentators claim that the unadjusted census is itself disruptive because it does not ensure certain groups of their rightful claims on political representation and Federal funding. These claims rely fundamentally on the conclusion that the adjusted counts better reflect the distribution of the population. As explained in the discussions of Guidelines One, Two and Three, the evidence supports the contrary conclusion.

Estrada asserted that the public good is better served by focusing on the potential benefits to millions of persons rather than on the limited number of Congresspersons and state legislators who would be affected by a decision to adjust.[138] As demonstrated previously, the evidence indicates that millions of Americans may be harmed rather than benefit from an adjustment. Moreover, we must remember that the Congresspersons and state legislators who would be affected by an adjustment are elected by and represent millions of Americans in the political process.

Comments by members of the public and by Estrada [139] noted that an adjustment would result in more equitable allocations of federal funding to states and cities, a consideration which in their view must be weighed against any disruptive consequences from adjustment. Again, this claim assumes that the adjustment provides a more accurate distribution of the population across states and localities, an assumption which is not warranted by the evidence.

Moreover, it has been demonstrated that an adjustment of the census would have very little effect on the distribution of Federal funds. The study in Appendix 15 [140] shows that less than one fifth of one percent of Federal funds would be reallocated as the result of an adjustment. Twenty-one or fewer states would receive additional funds from an adjustment. Fewer than half of all jurisdictions would be allocated additional funds as the result of an adjustment. As the study demonstrates, those jurisdictions that do benefit would receive on average only $56 in additional funds per "adjusted" person.

Thus, even if the claim that an adjustment would more accurately (and thus fairly) allocate federal funds were valid, the adjustment would not result in significant shifts of those funds.

Conclusion

Any decision will result in some level of disruption through legal challenges. On balance, the record indicates that a decision to adjust would likely be more disruptive than a decision not to adjust. A decision to adjust would clearly cause disruption in those States that have early redistricting deadlines. The assertion that persons are denied their rightful claims without an adjustment assumes that the distribution of the population is improved by an adjustment. Based on the evidence, this assumption is invalid. Thus, this guideline weighs in favor of a decision not to adjust.

Guideline Eight

The ability to articulate clearly the basis and implications of the decision whether or not to adjust shall be a factor in the decision. The general rationale for the decision will be clearly stated. The technical documentation lying behind the decision shall be in keeping with professional standards of the statistical community.

Explanation

It is the responsibility of the government to have its critical decisions understood by its citizens. We recognize, however, that the degree to which a decision can be understood cannot alone dictate an important policy decision.

The decennial census is a public ceremony in which all usual residents of the United States are required to participate. If the census count were statistically adjusted, the rationale for that action must be clearly stated and should be understandable to the general public. If the decision were made not to adjust, the elements of that decision must also be clearly stated in an understandable way. It will be the responsibility of the Department of Commerce and the Bureau of the Census to articulate the general rationale and implications of the decision in a way that is understandable to the general public.

This does not require the Bureau or the Department to explain in detail to the general public the complex statistical operations or inferences that could lead to a decision to adjust. But, as with any significant change in statistical policy, the government has the duty to explain to the public, in terms that most can understand, the reason for the change. If the decision is not to adjust, (that is not to change) the public will be informed as well.

The last part of the guideline ensures that the methods, assumptions, computer programs, and data used to prepare population estimates and adjustment factors will be fully documented.

The documentation will be sufficiently complete for an independent reviewer to reproduce the estimates. These standards apply to the post-enumeration survey estimates, the demographic analysis estimates, and the small area synthetic estimates.

Discussion

The general rationale for this decision is clearly stated in the first section of this report. The technical documentation underlying this decision is in keeping with the professional standards of the statistical community. Thus the Guideline has been satisfied.

However, the Guideline could have been met if the decision had been to adjust. The Census Bureau has done a laudable job of keeping the public informed of the progress of the post-enumeration survey and the progress towards the adjustment decision. There is no doubt that the process of adjustment is complex and the statistical details of the process are fully comprehended by only a few individuals. Although I am sympathetic with these arguments, this would not have been an impediment to an adjustment. The general rationale could

[137] Tarrance, pages 2–3.

[138] Estrada, page 24.

[139] Estrada, page 23.

[140] See appendix 15. Murray, Michael. "Census Adjustment and the Distribution of Federal Spending." U.S. Bureau of the Census, May, 1991. (hereafter Murray).

have been clearly articulated. As Estrada notes, the public perception of the census "head count" is far removed from the actual process, [141] yet the general rationale for the census is well understood.

Conclusion

The requirements for this Guideline have been met. This Guideline does not weigh in favor of a decision either way since the requirements of this Guideline could have been fully met if the decision had been to adjust.

SECTION 3—SUMMARIES AND EVALUATIONS OF THE RECOMMENDATIONS OF THE SPECIAL ADVISORY PANEL

In this section I summarize the individual recommendations of each of the members of the Special Advisory Panel appointed to advise me on this decision, and the joint recommendation offered by Drs. Ericksen, Estrada, Tukey, and Wolter. After each summary I evaluate each recommendation.

Recommendation of Eugene P. Ericksen

Summary of the Recommendation

Ericksen recommends an adjustment. His argument relies substantially on a report co-authored by himself, Estrada, Tukey, and Wolter. He argues as follows: An adjustment will reduce the substantial error in the census and will correct for the differential undercount. The Bureau produced a demonstrably inaccurate census enumeration which can be fixed by means of PES estimates. PES estimates have been demonstrated to be both accurate and statistically reliable by evaluation studies of the 1990 decennial census. The racial differential undercount has again been demonstrated in the census, and the PES can correct for this clear and important bias.[1]

On Guideline One, Ericksen reports from his jointly authored analysis and other analyses that it is clear the adjusted count has been shown to be more accurate than the original enumeration. In Ericksen's view there is little doubt that the original enumeration is inaccurate. He states that the Census Bureau reported 13 million erroneous enumerations, 19 million omissions, and a PES net undercount rate of 2.1%.[2]

Ericksen says the basic flaw of the original enumeration is that it uses a method "designed for the well educated, middle-class family with reliable mail service." He argues that the method does not work for "those who do not read well, who live doubled up in an apartment, who live in drug infested neighborhoods with high crime rates, and who only occasionally receive mail." The procedure had such well demonstrated flaws that the 4.7 million undercount, and the 4.4% demographically estimated differential, was not surprising.[3]

Ericksen states that the PES was successful. The interviewing quality was high, imputation was minimal, and the matching error was very small. The evaluation studies suggest that the total error in the PES was minor. Correlation bias suggested that the PES underestimated the undercount, if anything. "The only reasonable conclusion is that the adjusted count is more accurate than the unadjusted count." [4]

On Guideline Two, Ericksen states that the adjusted data are consistent, complete, and of sufficient quality to be used for all purposes and at all levels for which census data are used. He cites the jointly authored report.[5]

On Guideline Three, Ericksen finds that "under any reasonable basis of comparison, the PES-adjusted enumeration is more accurate than the unadjusted enumeration." [6] The PES estimates are robust with respect to evaluation strata, and the effect of the PES biases on population shares is negligible. The estimates for the states whose Congressional delegation size might be changed by an adjustment are stable.

On Guideline Four, Ericksen says it is difficult to comment because of the lack of evidence. He interprets the available evidence from a National Opinion Research Center (NORC) study to suggest that most Americans would like to have the most accurate census and will trust the experts to make it so.[7]

Ericksen has no expert opinion on Guideline Five but notes that Jefferson lamented the lack of accuracy in the first census.

On Guideline Six, Ericksen feels sufficient data are available to make the decision now. Sampling errors for local estimates are reasonably small, and the PES evaluation studies indicate that bias is small.

On Guideline Seven, Ericksen admits having little comment. As a scientist he feels it is better to use improved numbers when available than to rush ahead and make errors.

On Guideline Eight, Ericksen believes that the results can be explained, and the technical documentation is in keeping with professional standards.

Evaluation of Recommendation

I agree the census had an undercount. I also agree that the evaluation studies demonstrated that the PES was well done. I do not agree, however, that the PES has the ability to correct distributional error. The grounds for my disagreement have been documented in the discussion of Guideline One.

I agree that the adjusted count, if more accurate, has been shown to be more accurate in a numeric sense at the national level. I do not agree that the adjusted count is more accurate in the distributional sense at lower levels of disaggregation. In addition, the erroneous enumeration and omission figures cited are Census Bureau estimates, which vary according to definition.[8]

The census used a variety of methods, including mail-out/mail-back, list enumerate, and list leave to fit different lifestyles. Class membership, education level, and reliability of mail service may explain some, but not all, of the census coverage problems. Recall that the personal enumeration censuses of 1940, 1950, and 1960 had even higher estimated undercounts. Thus, I disagree with Ericksen's notion that the census was "designed for the well educated, middle-class family with reliable mail service."

I do not agree that successful PES operations imply that the statistical manipulation required to go from its data to 4,830,514 blocks in order to produce a better count is a routine, automatic operation. I disagree that PES data, which are informative about the census, *can* be used to change the census in ways that make it distributionally more accurate.

I do not agree that merely because the Census Bureau can produce data that completely duplicate enumeration tables, that those numbers are of sufficient quality to be substituted for the census enumeration.

[141] Estrada, page 24.

[1] Ericksen, p. 1.

[2] Ericksen, p 2.

[3] Ericksen, p. 2.

[4] Ericksen, p. 3.

[5] Ericksen, p. 3.

[6] Ericksen, p. 3.

[7] Ericksen, p. 3.

[8] The numbers used by Ericksen are estimates derived from all P-sample misses (19,171,290) and all E-sample "Erroneous Inclusions/Unmatchable." (13,154,630) While defensible, this is but one extreme definition. For example, it does not take into account the role of Census imputations. The matter of estimating these two components is a matter of disagreement among professionals. See, for example, the discussion in a Memorandum from Howard Hogan to Pete(r) Bounpane entitled "Gross Census Errors," July 2, 1991, Bureau of the Census on these issues. See the discussion of this issue under Guideline One above.

I agree that the PES adjusted enumeration may be more accurate numerically. I do not agree that it is distributionally more accurate. While the estimates are robust for evaluation strata, there is considerable doubt cast on their homogeneity with respect to post-strata relative to states.[9]

I appreciate Ericksen's comments on Guidelines Four and Seven, although I do not agree with them. I agree with his comments on Guideline Eight. I agree with his comments on Guideline Six, except that sufficiency of data in Guideline Six has nothing to do with substantive outcome, as Ericksen's comments about the size of sampling error would seem to imply.

Recommendation of Leobardo F. Estrada

Summary of the Recommendation

Estrada recommends in favor of an adjustment. Estrada first spells out a general rationale for his decision which is followed by an exposition of his reasoning for each guideline. Estrada relies on the paper co-authored by Ericksen, Estrada, Tukey, and Wolter.

Estrada's general rationale begins with the observation that the 1990 census is sufficiently flawed to require adjustment. In particular, the undercount rate increased from 1980, the census omitted the largest number of persons ever, historical undercount differentials between blacks and non-blacks persisted, and the black non-black differential actually increased from 1980 to 1990.[10]

Estrada states that the observed pattern of undercount is consistent with prior censuses. The Census Bureau efforts to overcome the undercount in the enumeration failed for a variety of reasons relating both to the character of the population and to the nature of the census operation itself. "While the Census Bureau was able to improve its internal management systems, the national dynamics that comprise the U.S. became more complex." [11]

Estrada argues that the differential undercount was the real cause for concern. He asserts that it occurred due to a number of problems in census processes. Flaws in the census operation included inaccurate mailing lists, non-delivery of census forms, a lower than expected mail return rate, inadequate interviewer and enumerator staffing levels, delay in district office closings, enumerator errors, enumeration by last resort, missing data, the inclusion of 2 million non-data defined persons in the count, lack of non-English language forms, processing errors, lost forms, race and ethnicity misclassifications, geocoding errors, and duplicate records.[12] District offices in the largest cities with the most heterogeneous populations suffered more from these flaws than others resulting in more last resort, close-out and non-data defined enumerations among non-Hispanic blacks and Hispanics.

Estrada states that the cumulative effect of all these problems is that the 1990 census needs adjustment.

Estrada describes the post-enumeration survey (PES) as a high quality process. He ascribes the high quality of the PES to, among other things, on-site listing of livable structures rather than reliance on mailing lists in sample blocks, experienced interviewers, a non-response rate of less than 1% and a proxy response rate of only 2.4%, relatively early interviewing to overcome the forgetting problem, successful tracking of the 8% of the PES who were movers, the successful evaluation program, and the fact of matching or resolving non-match cases for 98.3% of the 173,000 housing units surveyed.[13]

Estrada says that PES estimates of undercount follow known and expected patterns; *i.e.*, blacks higher than non-blacks, young males among minorities most often undercounted, the West division higher than other divisions, Hispanics highest rates of all. This attests to the "reasonableness" of PES undercount estimates and shows consistency with demographic analysis.[14]

Estrada claims the quality of the dual system estimates is sufficiently high to justify their use, according to the Hoaglin and Glickman sensitivity analysis among others.

Estrada says that adjustment methodologies improve the proportional distribution at all levels of census geography. He relies on the Tukey work and the work of other consultants that show that improvement at higher levels of geography improves shares at lower levels.

These conclusions by Estrada end the general rationale section of his recommendation The remainder of Estrada's recommendation focuses on each guideline.

On Guideline One, Estrada begins by reviewing the results of the Census Bureau evaluations of the PES, the so-called P-studies. The missing data studies (P1, P2 and P3) show that the rates of noninterview are low and the imputation for the primary population items was also low. Alternative means of imputing missing data did not affect post-strata. A Special Advisory Panel (SAP) analysis shows that post-stratum shares are minimally affected by eight alternative ways of handling missing data, with one exception. Given the small number of imputations required for the PES, alternative methods would have small effects on the outcomes.

Estrada says that the matching error studies (P7 and P8) confirmed that the high quality of clerical matching and matching of movers was performed successfully.

Estrada writes that the correlation bias studies (P13, P14, and P17) show strong correlation bias in the PES. Although for some this casts doubt on the dual system estimates, for him there is another side to the coin—"the undercount would be underestimated, particularly for minority populations. Whether the underestimation of undercount caused by correlation bias balances the biases toward overestimation of the undercount caused by missing data needs to be investigated, but the chances are they offset each other." [15]

Estrada states that other studies of data quality from the PES (P4, P5, P5A, and P6) show that the PES was not seriously impaired by problems of the quality in the reported census day addresses or fabrication.

Estrada says that those studies related to erroneous enumerations (P9, P9A, P10) show that erroneous enumerations were concentrated in particular evaluation post-strata. The census had higher rates of erroneous enumerations in minority areas and rural areas. Some significant changes would have occurred had matching of cases reported as erroneous enumerations been done by expert matching. On the census side there was a low error rate in matching, but more detailed analysis indicates that erroneous enumerations due to matching were more likely in two evaluation post-strata—non-minority areas outside the central city in the Northeast and West.[16]

Estrada claims that the study on late-late census enumerations (P18) shows that the addition of these data had an insignificant effect on the undercount rate. Similarly, balancing error was not a problem.

[9] See P12, and the discussion of Guideline Three above.

[10] Estrada, page 2.

[11] Estrada, page 3.

[12] Estrada, pages 4-6.

[13] Estrada, pages 6-8.

[14] Estrada, pages 8-9.

[15] Estrada, page 14.

[16] Estrada, page 16.

Estrada believes that the total error model (P16) indicates that errors introduced in the PES were small and tended to equalize racial differentials in undercount.

On Guideline Two, Estrada states that a strong argument can be made that the requirement for local area accuracy can be satisfied by showing that adjusted counts are an improvement on the average for the principal uses of census counts. He claims it is appropriate to judge the adjusted counts at higher levels of aggregation than the block.

Estrada acknowledges that the Census Bureau study on heterogeneity (P12) shows mixed results with respect to the homogeneity assumption with respect to poststrata. "The research 'flags' the need to be aware of State effects [overwhelming poststrata effects]." [17]

On Guideline Three, Estrada acknowledges that the Census Bureau study on coefficients of variation (P15) showed that estimates of variances and covariances for smoothed and unsmoothed adjustment factors were larger than expected. However, he cites the Hoaglin and Glickman study as demonstrating the robustness and stability of the dual system estimators under different statistical treatments.

On Guideline Four, Estrada argues that if the Secretary adjusts using the best tools available, the reputation of the Census Bureau will be enhanced. The census process must incorporate adjustment as its final step. Estrada interprets the National Opinion Research Center (NORC) poll as indicating that the decision to adjust is slightly more likely to improve participation in future Censuses.

On Guideline Five, Estrada states that the innovation of adjustment is in keeping with prior Census Bureau efforts to meet the intent and spirit of the Constitution. The courts have already held that adjustment can be Constitutional.

On Guideline Six, Estrada states that "all the proposed studies have been completed, the data tables made available and the Census Bureau has had sufficient time to fulfill the concerns set out by [this guideline] in time for the Secretary of Commerce to make his decision." [18]

On Guideline Seven, Estrada states "Without denying the fact that there are State officials who feel imposed upon and elected officials (and potential challengers) who suffer from uncertainty as to when the boundaries of their districts will be 'fixed,' the actual consequences [of the census being adjusted and these figures not being available until July 15, 1991] are that a couple of Congressional seats will shift from one State to another; that delays will occur in redistricting, and that edges of many recently designed districts will have to be slightly modified to meet the one-person, one-vote requirements." [19]

Estrada says these disruptive consequences must be weighed against the fact that a census adjusted for deficiencies will provide a more equitable allocation of persons to each district, and a more equitable allocation for all other census purposes. The public good is better served by focusing on the potential benefits of adjusting the census to millions of persons rather than on the limited number of Congressmen and Congresswomen and legislative officials who will be affected by the July 15, 1991, decision to adjust the census and the subsequent release of adjusted numbers. [20]

On Guideline Eight, Estrada states there is an implicit assumption that the public understands the standard census methodology. However, their perception of what the census is—is far from the real census. Thus, both the real census and the reason for adjusting the census must be understood by the public. The public must understand the context of the PES in the census process. An informed public will accept the need to adjust if provided with concepts to understand the logic of the method. [21]

In conclusion, Estrada notes that the census has suffered from a persistent differential undercount. The evidence overwhelmingly demonstrates that the census count can be improved by adjustment. The PES adjustment factors have an advantage over demographic analysis in providing more specificity about the undercount. Adjusted counts will be more equitable and assure equal representation. Therefore, the Secretary of Commerce should adjust the census.

Evaluation of Recommendation

I do not agree that it follows that even were the 1990 census sufficiently flawed to require an adjustment, an adjustment is possible. The facts cited comparing 1980 and 1990 are a necessary, but not sufficient, grounds for considering an adjustment. A methodology must be available that will achieve a successful distributional correction.

I agree that the differential undercount is regrettable, and a cause for serious concern. I do not agree with Estrada that the flaws cited in the census are tied directly to that undercount. I agree that no matter how the differential came about, one would want to fix it *if one could*.

I agree that the PES was successful. However, I do not agree that the PES estimates followed all expected patterns. For example, in the discussion of Guideline One, above, serious questions are raised about its success in finding black males, and its "over compensation" for older females. In fact, PES results are frequently inconsistent with demographic analysis. [22]

I believe that the Hoaglin and Glickman study can be interpreted to show not robustness, as Estrada says, but that it can be interpreted to show that thirteen different models produce thirteen different sets of adjusted counts. These counts may have been close to one another, but not necessarily be an improvement over the census. Furthermore, as I noted in the discussion of Guideline Two there are other sources of variation due to statistical modeling.

I do not agree that the conclusions reached with respect to the Panel correlation bias studies are as clear as Estrada asserts. As Special Advisory Panel member Wachter suggests, the undercount may be underestimated by correlation bias effects not because of differential misses, but by differential erroneous enumeration rates when holding misses constant. [23]

I believe that Estrada's discussion of erroneous enumerations reaches the opposite conclusion from what the studies find: Differential erroneous enumeration rates by evaluation post-strata are a cause of concern, because they leave open the real possibility of differences between processing offices in how well the PES was carried out.

I agree that the total error model is experimental, but I disagree that the expression "total" is appropriate. Not all errors are included in it, only those errors that could be estimated on the basis of the PES. While the study of total error is encouraging, it is not yet dispositive with respect to the utility of the model.

Estrada acknowledges that P12 shows mixed results with respect to heterogeneity of post-strata. Thus, his assertion that requirements for local area accuracy are satisfied by "average improvement," and that only higher than block levels of aggregation need be considered, seems to me to contradict his acknowledging that local area

[17] Estrada, page 19.

[18] Estrada, page 23.

[19] Estrada, page 23.

[20] Estrada, page 23–24.

[21] Estrada, pages 24–25.

[22] See the discussion of Guideline One above.

[23] Wachter, pages 12–13.

accuracy needs to be satisfied. In fact, heterogeneity at the block level would mean that Guideline Two has not been satisfied.

As noted earlier, I believe that the Hoaglin and Glickman study can be interpreted as demonstrating a clear lack of robustness: Since accuracy at the block level is the goal, a process that allows thirteen different models to produce thirteen different estimates that differ only a little from one another, is not adequate. Differing a little at the high level of aggregation of the Hoaglin and Glickman work may mean differing dramatically at the block level.

I do not agree with Estrada's comments on Guideline Seven. The adjustment, as envisioned, will, in fact, not provide a more equitable allocation of persons to districts as he assumes. In my opinion, the lack of distributional accuracy is precisely why the adjustment is flawed as a correction for the census counts.

I do not agree that adjusted counts will be more equitable as Estrada claims in his discussion of Guideline Eight. In fact, they will not be more equitable distributionally, which is the criterion for determining whether an adjustment would improve the accuracy of the counts.

Recommendation of William Kruskal

Summary of Recommendation

Kruskal recommends against an adjustment. He uses the word "modification" rather than adjustment since the latter term suggests to him that "we really know how to improve the Census enumeration." [24] The primary reason for recommending against adjustment is that "we do not know with any confidence how to make such improvements . . . and we will not know in a relevant time scale." [25] Although "the proposed modifications are clever and technically interesting, the method turns on highly specialized assumptions and we simply do not know how robust the output results are against realistic errors in those assumptions." [26] The proposed modifications are complex, impossible to explain clearly for a general audience and their use is "likely to increase already existing apprehensions about manipulation and big brotherism in Washington." [27] The modified estimates might well introduce more error than they clear up, without anyone being aware of such an imbalance.

On Guideline One, Kruskal contends that there is no conclusive evidence that the modification removes more error than it introduces, and does not expect any convincing arguments anytime soon. The major gap in assessing comparative accuracy is the uncertainty about the "capture-recapture" model.[28] The implicit assumption of uniform capture probability is the most troublesome. Knowledge about the degree of output error caused by the non-factuality of this assumption "is just what we do not have, indeed cannot have, for the post-enumeration process." [29]

Later Kruskal notes that Guideline One calls for the highest professional judgment from the Census Bureau. "The highest level of professional judgment requires vigorous argument and discussion not only within the Bureau but in groups made up both of Bureau and outside statisticians and others. That vigorous and public discussion we have not had in nearly adequate amount." [30]

On Guideline Two, Kruskal's only comment is that synthetic adjustment is based on a simplifying assumption that is known to be wrong, which in turn throws great weight on the calculations of stability, given reasonable error structures.[31]

On Guideline Three, Kruskal's impression is that "choice of the so-called smoothing procedures was profoundly based on post-enumeration survey (PES) results," [32] which is not in keeping with the guideline. He questions whether "that in *major* respects the choice of procedure was made before the PES results were in hand," but time did not permit a full investigation on his part.

On Guideline Four, Kruskal feels the extraordinarily complicated procedures will undercut public confidence in the census. On Guideline Five, Kruskal has no comment. On Guideline Six, Kruskal believes that "timely data and analysis are not really at hand." [33] On Guideline Seven, Kruskal does not see how "this cuts in the present context." [34] On Guideline Eight, public explanation will be difficult because of the complexity and the choice of one of many such methods available.

Kruskal notes that the Guidelines "tilt against modification," but "that is hardly novel." [35]

Without resting his views solely on the guidelines, and instead on his "partly formulated and internalized professional criteria, along with [his] internalized civic standards," [36] Kruskal still recommends against an adjustment. He expresses concern about the large numbers of estimated counts and the inherent problem of putting together the millions of estimated differences between the count and the adjustment. He closes by noting that modifications that increase counts can, in fact, harm, by moving the proportions of the population in a given area in the wrong direction.

Evaluation of Recommendation

I agree that the census modifications lack robustness. Thus, Kruskal does not interpret the Hoaglin and Glickman studies as do plaintiffs' panel members. He recognizes that the adjustment may introduce more error than they correct without anyone knowing it.

I agree with Kruskal's criticism of the "capture-recapture" model upon which the DSE is based. He notes, in particular, that its assumption of uniform capture probability is most troublesome.

I agree with Kruskal's belief that there has not been an adequate vigorous and public discussion of the merits of adjustment. However, I disagree with his statement that the lack of such a discussion means we are not able to determine whether Guideline One is adequately met.

I disagree with Kruskal that, in terms of Guideline Three, there was no prespecification. He asserts that smoothing procedures were based on PES results. His comments implies a standard that would, in Guideline Three terms, preclude ever meeting prespecification requirements.

I agree with his comment that increasing counts can move proportions of the population in a given area in the wrong direction. That comment means that he, too, is concerned with the problem of distributive accuracy, and that he shares a concern about whether the proposed procedures deal with it adequately.

Recommendation of Michael McGehee

Summary of Recommendation

McGehee strongly recommend[s] that no adjustment be made to the census. There is no compelling evidence that suggests that the

[24] Kruskal, page 1.
[25] Kruskal, page 1.
[26] Kruskal, page 1.
[27] Kruskal, page 1.
[28] Kruskal, page 2.
[29] Kruskal, page 3.
[30] Kruskal, page 5.
[31] Kruskal, page 3.
[32] Kruskal, page 4.
[33] Kruskal, page 5.
[34] Kruskal, page 5.
[35] Kruskal, page 5.
[36] Kruskal, page 6.

PES [post-enumeration survey] will provide estimates that are any closer to the true population totals for the eight million blocks [illegible] United States. Indeed, there is [illegible]ant evidence to suggest that [illegible]ment will move the population of many blocks further away from their true populations.[37]

Persons have always been missed in the census for a variety of reasons. Statistical adjustment is the most recent proposal to address the situation.[38]

McGehee states that adjustment numbers are estimates just like census counts: there is no way to determine which is closer to the true population, other than assumption and judgment. The evaluations of PES data "rested on pre-conceived assumptions of how the data would appear."[39] The results often fell outside the limits predicted from these assumptions. Rather than accepting the conclusion that the process is flawed, the assumptions were modified. He has no confidence in this reasoning. He refers to the problem in computing margins of error (variances) for local estimates as an example of this problem. "It is a strong indictment of the entire process, however, when evaluation procedures are not clearly understood by those using them * * *. The entire process has tended to produce more, rather than less, uncertainty."[40]

McGehee gives, as an example of the [illegible]inty created, the large difference [illegible]uction matching effectiveness ra[illegible] between Albany and Kansas City (87.20% v. 93.49%). Why this discrepancy exists is unknown and "no documented evidence can be presented which clearly explains this problem."[41] Adjustment proponents will argue that in the aggregate these problems are small and thus "the differences at lower levels should be overlooked because they become insignificant at the aggregate level."[42] McGehee disagrees, pointing to Guideline Two requiring accuracy across all jurisdictional levels. Furthermore, variation at the aggregate level, McGehee contends, is discounted by proponents by modifying the assumptions upon which the conclusions have been based.

"Decisions made during the DSE process, and the assumptions on which they stand, dramatically alter the adjustment results. A politically 'better' count cannot be defended if it is shown that the assumptions on which it rests are changeable."[43] Because of the widespread use of census figures, they must be defensible. The Bureau has maintained public confidence in its numbers over the years by "its meticulous approach to detail and its dogged adherence to maintaining the quality of Bureau data as the true standard."[44] Adjustment will undermine the public's confidence in this track record. A decision to adjust should be treated as political, and be forced to undergo the same Congressional scrutiny as other such decisions.

McGehee continues his argument by discussing the capture-recapture methodology. He uses an analogy to compare the PES to counting bears in a game preserve. He notes that the heterogeneity in game wardens' background and abilities, in the types of bears and their physical characteristics and in the terrain will lead to differences in how well the bears are counted. In similar ways, the enumerators' characteristics, the characteristics of the population the enumerator is counting and the environment in which the enumerator is working will all have effects on the outcome of the PES. These problems are compounded by the fact that PES records must be matched back to the census and the ability of matchers may be heterogeneous.[45] To identify the weight given to each of these variables, regression models are used to determine their individual effect. How these regression models are specified in the PES process is constantly changing. How to combine these variables into a larger number and how to compare various strata are issues of judgment on which individuals may differ.[46]

McGehee says that comparisons of data to the "correct" or "true" population are often made. The "correct" population is derived from a series of assumptions and thus results are simply theories. After reviewing the data, it is clear that the proposed adjustment does not meet the criterion of being usable across all jurisdictional levels nor is it robust at local levels to reasonable alternatives. The idea of using the PES to adjust the census is so complicated and so subjective, that no reasonable person can agree that it should be contemplated or that the process will be explicable to the general public.[47]

McGehee next turns to the issue of comparing the accuracy of the PES to the Census. Matching PES and census records is the key to assessing the relative success of the PES and the census in counting people. His "analysis shows that the PES fails to demonstrate a 'better' record of counting people than the Census. Indeed in many instances it cannot demonstrate that it did *as well* as the Census."[48] In support of his assertion McGehee presents a cross tabulation of census match codes by race and ethnic origin. He also does so for the PES. Although "time does not permit extensive analysis of this data,"[49] he does note that twice as many Hispanics in the census left the race question blank as in the PES. More Hispanics identified themselves in the category "other" in the PES than in the census. "On a superficial basis, the results raise very significant questions whether adjustment will, in fact, yield greater accuracy than the census."[50]

McGehee states that the rationale for using the PES to correct the differential undercount rests on the assumption that as the black population increases in each block cluster, the PES will do a better job than the census in counting people.[51] It is appropriate then to compare the "best" and "worst" census and PES numbers within each block cluster and see how these comparisons change as the concentration of blacks increase over clusters.

McGehee argues that since errors occur in both the census enumeration and the PES survey, judgments had to be made as to whether it was correct to include them. These judgments are critical in determining the success or failure of the PES or the census. In those cases where judgments were made, one can get a range of estimates of quality by assuming that all judgments should have gone in favor of omission and, alternatively, all judgments should have gone in favor of inclusion.[52] Best and worst confidence level scenarios for the census and the PES in each block cluster are carried out. These comparisons are displayed by ranking the results in order of the proportion of blacks in the cluster, since research indicates "that as the percentage of black population within a cluster increases, the effectiveness of census coverage decreases."[53]

[37] McGehee, page 8, emphasis in the original.
[38] McGehee, page 2.
[39] McGehee, page 3.
[40] McGehee, page 4.
[41] McGehee, page 4.
[42] [illegible]hee, page 5.
[43] McGehee, page 5.
[44] McGehee, page 6.
[45] McGehee, pages 8–10.
[46] McGehee, page 11.
[47] McGehee, page 12.
[48] McGehee, page 14.
[49] McGehee, page 19.
[50] McGehee, page 19.
[51] McGehee, page 20.
[52] McGehee, page 21.
[53] McGehee, page 25.

McGehee uses six graphs to present these results. "When comparing the best census scenario with the worst PES scenario one sees that the census does a dramatically better job of correctly counting people than the PES. ... What is surprising, however, is the potentially dramatic performance shown by the census in those clusters where the black population is between 50% and 75%. Even more surprising is the very close correlation between the census and the PES in clusters where the black population is greater than 80%. In fact, the Census has a higher confidence level than the PES in those clusters where the black population is between 80% and 85%. This flies in the face—and graphically demonstrates the fallacy—of the argument put forward by the proponents of adjustment." [54] The PES does not necessarily outperform the census. Even if one accepts the midpoint between the best and worst PES results, the census exceeds this level and the PES does not outperform the census in clusters containing a large number of blacks.[55]

McGehee then turns to the guidelines. In his discussion of Guideline One, he finds the entire concept of adjustment on "the outer limits of statistical research." [56] The assumptions underlying the evaluations of the PES are so arbitrary and fluid that little weight can be attached to their assessments of PES quality. Therefore, Guideline One cannot be met since one cannot prove that the PES is better than the census.

On Guideline Two, he notes that variances between processing offices and evaluation strata are outside expected levels and at the district office level there was such variation it could not be reconciled. Adjusted numbers are inconsistent at the State, city, and subcounty level and suffer from serious quality concerns.[57]

On Guideline Three, McGehee asserts that the adjusted counts have not been shown to be more accurate than the census enumeration. The determination of quality is dependent on many assumptions and judgments.

McGehee says that the manipulation of assumptions in evaluation studies undermines confidence in all ongoing statistical data collection and therefore Guideline Four cannot be met.[58]

[54] McGehee, page 28.
[55] McGehee, page 29.
[56] McGehee, page 31.
[57] McGehee, page 32.
[58] McGehee, page 33.

McGehee claims there remain legality questions about adjustment that need to be answered with respect to Guideline Five. On Guideline Six, McGehee states that sufficient data are available to suggest that the PES was flawed and the analysis of the data is insufficient to justify a decision to adjust the census.[59]

On Guideline Seven, McGehee finds that the mere fact of a possible adjustment has caused consternation and difficulties in state legislatures. The lack of consensus on the desirability and statistical feasibility of adjustment will result in extensive legal battles.[60]

Finally on Guideline Eight, McGehee asserts that the entire process is so complicated and difficult to understand, even by professionals, that a general rationale cannot be clearly justified. To the degree that the process is explained successfully people will become aware of the kind of manipulations underlying it and the integrity of the statistical process will be forever compromised. Adjustment is to correct an inequity, which is not a statistical problem but a political and societal problem that should be dealt with by the Congress.[61]

Evaluation of Recommendation

I agree with McGehee that the results of the PES fell outside expectations. The error variance around local estimates are an example of this problem.

I agree with McGehee's citing large differences in production matching effectiveness between processing offices as indicators of uncertainty rampant in the PES data. However, evaluation studies of the PES have not found the kind of systematic effect alleged.

I disagree that the link between the Bureau's credibility and its aversion to schemes that tend to devalue the census itself is a reason for avoiding adjustment.

I agree with McGehee's criticisms of the capture-recapture methodology. He raise issues not brought out elsewhere that cast doubt on its validity for use on human problems. I agree with his notion that characteristics of interviewer, interviewee, and setting interact to affect the quality of information, and find McGehee to persuasively elaborate the idea. I believe that McGehee's ideas support criticisms of Kruskal and others that the method is flawed fundamentally.

I disagree that if an adjustment were made it would not be explainable to the public. Since the decision not to adjust is just as complicated, this statement does not seem to have merit as an argument against adjustment.

[59] McGehee, page 33.
[60] McGehee, page 34.
[61] McGehee, page 35.

Although I concluded that an adjustment would degrade the quality of the population distribution as compared to the census, I do not agree with McGehee's explanation of why the PES did not do as well as the census. He presents an analysis showing that, in a sample of block clusters, as the percentage of blacks within a cluster increases, the census actually performs better than expected. McGehee claims that this analysis casts serious doubt on the argument that *ipso facto* a PES based adjustment will necessarily reduce the differential undercount of blacks. I find his argument at best anecdotal and not compelling.

I agree with McGehee's conclusions that, on the basis of his analyses, arguments for adjustment based on Guidelines One, Two, Three, and Six are not adequate: The census remains more accurate than the PES; adjusted numbers are inconsistent at different levels of geography, and the quality of the PES is too dependent on assumptions, not facts and analysis.

McGehee argues on Guideline Seven that disruption is already occurring. This argument lacks support. He cites no evidence that adjusting or not adjusting will differentially contribute to disruption. Thus, I find that his arguments that this Guideline argues against adjusting are not relevant.

I disagree with his belief that the technicalities cannot be explained. Rather, I note that the process has been open, the Bureau has gone to great lengths to document its activities, so that there was no lack of ability to explain adjustment.

Recommendation of V. Lance Tarrance, Jr.

Summary of Recommendation

Tarrance recommends against an adjustment. He has chosen to concentrate on the public policy implications of a decision, not only because it is his area of expertise but also because he is "convinced that the impact of changes to the enumeration totals on the operations of our government—at the federal, state and local levels—would be disastrous." [62] Tarrance's lengthy introductory remarks are followed by a discussion of the guidelines.

Tarrance states that the perception that if the Bureau discovers how many persons it missed it should be an easy task to correct census results is

[62] Tarrance, page 1.

incorrect. In fact, there is no consensus on how to fix the counts among statisticians or other experts. Two polls—March 1990 and April show no consensus on including estimates of missed persons in the count. Whites were evenly split; non-whites preferred a synthetic adjustment.[63]

Tarrance says that more important than the statistical quality of the numbers is the public policy aspects of an adjustment. These include "the paralyzing difficulties that changing the numbers will cause in accomplishing redistricting . . . for all levels of the electoral system; the damaging perceptions that will be given to the public about the two different sets of numbers from the census; the troubling uncertainties surrounding even statistically acceptable numbers . . ." [64] Such policy difficulties should not be dismissed as many proponents of adjustment have done.

Tarrance asserts that lost in the debate fostered by adjustment advocates are the following points of decisive importance: (1) The adjustment process is complex, not well understood, without precedent and evaluations of it are judgmental; (2) synthetic estimates below the State level will never be more accurate than census counts; (3) the deadline of July 15, 1991, has not allowed enough time for adequate evaluation of the adjustment process or its product; (4) two sets of numbers may create "chaos" for the 1992 elections; (5) the trust in census confidentiality and the belief in the need to cooperate with the census will be further eroded; (6) resources may be denied to future census activities because "adjustment will take care of all problems" will be the expedient prevailing attitude; and (7) accepting adjustment will invite " 'inside manipulation' of numbers for political purposes." [65]

Tarrance says that "The adjustment process being used can produce an array of different results depending on the choice of assumptions and/or statistical methods employed. . . ." [66] Thus, the issue is not technical, but judgmental, as the decision calls for an assessment of the consequences of a decision. Whatever the decision, litigation will ensue, but a decision against adjustment "may be the beginning of a more reasoned look at the problem." [67] The Constitution says Congress shall determine how the census is to be conducted; therefore Congress should settle this issue, if at all possible, rather than the courts.

Tarrance quotes a statement made by co-chair Ericksen in 1980: "The undercount adjustment procedure needs to be statistically sound and politically credible," and goes on to state that the controversy has increased, in fact, and Ericksen's 1980 position is even more compelling today. Given the confusion and possibly paralyzing effects of adjustment, the best solution is not to adjust the census today, but to consider the proposal to adjust intercensal estimates as is done in Australia, Finland, and Spain.

On Guideline One, Tarrance first notes that statistical sampling only produces accurate results when sample sizes are sufficiently large, and for small jurisdictions this is simply not the case. Some small area counts will be made less accurate by an adjustment and the question is how we deal with such areas. There are a host of questions about tradeoffs among communities in accuracy that remain unanswered.

Furthermore, he points out that accuracy is a point of fundamental definitional differences between law and statistics: law needs certainty, statistics accepts a range of uncertainty about numbers it still considers accurate. "Any court settlement directing adjustment will *necessarily* require the arbitrary choice of numbers which have been derived from methods that statisticians would ordinarily hedge about. . . . It is paradoxical that those same interests who are faulting the Bureau of the Census for not having counted all persons are at the same time putting inordinate trust in that same agency to transcend the limits of statistical 'estimating!' " [68]

Tarrance argues that:

> The important fact that is buried in the mass of rhetoric about the need to correct inequities resulting from undercounting is that the numbers will undoubtedly be less accurate for many areas below the state level. The reality is that the adjustment process will not find those persons who were missed by the original enumeration and include them where they were not counted before. . . . Some correctly counted blocks could have persons added to their count; some correctly counted blocks could have persons deleted from the census count, and incorrectly counted blocks might not have any changes made to their numbers.[69] In addition, the post-enumeration survey (PES) is not able to handle all forms of counting errors with equal adequacy. Thus, misallocation can still occur even with adjusted numbers. Ultimately, "the final numbers are chosen from a range of possibilities that are dependent upon the choice of assumptions; there is a great deal of 'inside' judgment involved, and although [he has] no reason to doubt the experts at the Bureau of the Census who have had to make the hard choices, it is still troublesome that there is an opportunity for different results to be obtained by the use of different methods or assumptions." [70]

On Guideline Two, Tarrance states that a lack of usability for redistricting is a major deterrent to proceeding with adjustment, because of the conflicts having two sets of numbers will generate. "The realities of redistricting at the state and local level, combined with the possibilities for endless litigation, are such that it would be naive to believe that synthetic numbers will be usable . . . for the purposes of redistricting and reapportionment." [71] With two sets of numbers, redistricting plans will likely end up in court and the likelihood of "chaos" for the 1992 elections seems ever more probable.

On Guideline Three, Tarrance is most troubled by "the acknowledged fact that different methods using different assumptions produce different results." [72] As an example he notes that small numerical differences lead to large consequences in reapportionment and redistricting. "It is all too obvious that the procedures being used will not produce robust numbers and that it would be possible to obtain an array of population counts which could have very different effects upon apportionment." [73]

The requirement for pre-specification in Guideline Three concerns Tarrance, as some procedures were prespecified and some were not. In particular the decision not to combine demographic analysis with the PES was made by staff, in stream. This is an example of an attitude of "if the numbers don't come out the way we think they should, we can change plans" which is "diametrically opposed to what good government policy should allow. Furthermore it is clear that the adjustment process is a statistical operation which has never been done before and there are many last-minute decisions being made." [74] Tarrance

[63] Tarrance, page 2 and Appendices.

[64] Tarrance, pages 2–3.

[65] Tarrance, pages 4–5.

[66] Tarrance, page 6

[67] Tarrance, page 7.

[68] Tarrance, page 13.

[69] Tarrance, page 13, emphasis in the original.

[70] Tarrance, page 16, emphasis in the original.

[71] Tarrance, page 18.

[72] Tarrance, page 19.

[73] Tarrance, page 19.

[74] Tarrance, page 21.

expressed uneasiness that "special interest pressure to adjust was pushing an incompletely researched or insufficiently tested statistical operation to a very shaky end".[75]

On Guideline Four, Tarrance states that a decision to adjust would have a far-reaching impact on future census efforts. Future censuses might be adversely affected as the Congress might well cut census funds, using the logic that an adjustment will fix the count anyway. Mayors and other local officials would question the necessity for their efforts on behalf of the census. The adjustment controversy might very well erode the already tenuous confidence of the public in the Census Bureau. The controversy surrounding the count should lead to imaginative ways to take the census in 2000, such as rolling samples, the "bare bones" head count, *etc.;* and legislative proposals immediately after the adjustment decision.

On Guideline Five, Tarrance states that Congress should determine how the census is to be conducted as required by the Constitution. Congress could also direct program solutions to resource allocation inequities.

On Guideline Six, Tarrance is convinced that the entire process has been rushed in an attempt to meet an arbitrary deadline. There has not been enough time for the evaluations. Given the controversy and that a general consensus has not developed, the adjustment should not be done without "the most exhaustive study and analysis of the data," which there has not been enough time to do.[76]

On Guideline Seven, Tarrance notes that the Special Advisory Panel met with representatives of the National Conference of State Legislatures. Technicians who must do the redistricting believe that they will be "paralyzed" by the "endless litigation" two sets of numbers will provoke if the census is adjusted,[77] although the very existence of two sets of numbers may be problematic. An adjustment would be most threatening to the creation of redrawn electoral districts for the 1992 elections.

Adjustment, according to Tarrance, will set a precedent for adjusting future censuses. He notes that one person miscounted in the PES represents from 500 to 1,000 persons that would be added or subtracted to develop adjusted numbers. The opportunity for, or perception of, manipulation to achieve desired ends will remain, but once adjustment is routine and not subjected to the scrutiny that it is now, the rigor of public examination to assure that manipulation does not occur will wane, and the risk, therefore increase.[78]

On Guideline Eight, Tarrance states that few people, even expert statisticians, really understand the process being used. He offers several examples of procedures and results of adjustment that are not well understood and states that it is impossible to articulate the complicated statistical procedures to the average person.

Evaluation of Recommendation

I disagree with the implication of Tarrance's discussion of public policy considerations that results of polls should play a substantial role for or against adjustment. I also disagree that if there is consensus that a particular adjustment would improve the counts and consensus that the adjusted counts are better than the enumeration, then an adjustment could be done based solely on that consideration.

I agree with Tarrance's point that there is support for not adjusting because of disruptive consequences for redistricting efforts.

I agree with Tarrance that the seven points of importance he cites, *i.e.,* complexity, lack of accuracy of synthetic estimates, inadequate time for evaluation, two sets of numbers leading to "chaos" for 1992, erosion of trust in census confidentiality, adverse consequences for funding future censuses, and the danger of inside manipulation, are valid expressions of concerns affecting the application of Guidelines One, Three, Six, Seven, and Eight.

I agree with Tarrance's discussions of lack of robustness which occur throughout the discussion. The point is made by him that judgment plays a substantial role in the choice of adjustment procedures. This is a flaw in the adjustment process pointed out in the discussion of Guideline Three, above.

I agree that Guideline One's requirements for accuracy are not met. The problem of misallocating people—even if one counts them correctly at a "higher" geographic level, is raised and documented. I agree that the arbitrariness of outcomes depending upon choice of assumptions is a fundamental weakness of adjustment.

I disagree that two sets of numbers will cause sufficient chaos to make either set not "usable" in Guideline Two terms. This is not the definition of usability intended by the guidelines. In fact, the effects of the numbers, if accurate and usable to the block level, should not play a role in the adjustment decision with respect to Guideline Two. This argument does not raise a bar to adjustment.

I agree that prespecification may be a cause for concern. However, because the prespecifications, such as the decision not to combine demographic analysis and PES results, were professionally done by career Census Bureau staff, I find that they impose no bar to adjustment according to this guideline.

I agree with Tarrance's assertion that adjustment will have an adverse effect on future censuses.

I do not agree that there has not been enough time for the PES evaluations.

I agree with the evidence as cited, including a meeting by the SAP members with representatives of the National Conference of State Legislatures, supporting, anecdotally, a prediction of endless litigation to be engendered by two sets of numbers, if an adjustment is made. I agree that there will be an increasing risk of future manipulation of the counts through adjustments if the precedent is set. This point is acknowledged in the discussion of Guideline Seven, above.

I disagree that the adjustment cannot be explained adequately, should it occur. I believe there is sufficient documentation to do so. I disagree with Tarrance's interpretation of the role of Guideline Eight on this matter.

Recommendation of John W. Tukey

Summary of the Recommendation

Tukey recommends an adjustment. He relies on the same report submitted by and coauthored by Ericksen *et al.* He argues that each and every one of the technical Guidelines are supportive of adjustment and the key Guidelines One and Three are indicative of an adjustment.[79] Tukey addresses the guidelines in the order given here.

On Guideline Four, Tukey states that a decision to adjust will enhance the Bureau's reputation and facilitate future operations, while a decision not to adjust may hinder future census efforts.

Tukey states that the questions raised in Guideline Five have been before the courts several times, and all decisions rendered permit adjustment.

On Guideline Seven, Tukey states that the Guideline must refer to aspects of orderly transfer of political representation that could not be

[75] Tarrance, page 22.

[76] Tarrance, page 27.

[77] Tarrance, page 28.

[78] Tarrance, page 29.

[79] Tukey, page 1.

representation that could not be
Tarrance, page 27.

anticipated in March 1990. There are no aspects.

Guideline Eight. Tukey states that uideline can easily be met. The technical documentation lying behind the adjustment decision is in keeping with the professional standards of the statistical community.

On Guideline One, in Tukey's professional judgment, the adjustments based on the post-enumeration survey [PES] have been prepared based on the highest professional judgment, and are more accurate, both as to numbers and as to shares, than the raw original enumeration.

On Guideline Two, Tukey notes that, since the Bureau is preparing consistent and complete counts down to the block level, there is "no bar to adjustment." [60]

On Guideline Three, Tukey says that the Bureau has stuck to prespecified procedures. Dr. Robert Fay and consultants Drs. David Hoaglin and Mark Glickman have done a series of studies testing different statistical models that agree with one another and have proved to be good.

On Guideline Six, Tukey states there should be no questions raised about nonadjustment because of inadequate data by 15 July 1991.

Tukey ends with a post-script that that the existence of sensitivity of tment to reasonable choices should o bar to adjustment, as long as it is small. The single prespecified procedure produces small sampling errors in comparison with post-stratum to post-stratum differences in adjustment factors to make it clear that adjustment provides smaller combined error than non-adjustment.

Evaluation of Recommendation

I disagree with the assertion that a decision to adjust will enhance the Bureau's reputation or facilitate future census efforts. In fact, other SAP members assert the opposite. [61]

I agree that Guideline Five is not a bar to a decision to adjust.

Tukey's interpretation of Guideline Seven, while unique, would not change the role this Guideline plays in the adjustment decision.

I agree Guideline Eight can be met.

I disagree that the analysis of Guideline One indicates that the Guideline has been met with respect to shares. Since the adjustment must clearly be shown to be superior to the census, controversy over this very important role played by census numbers indicates that the Guideline has, in fact, not been met. [62]

I disagree with Tukey's argument that Guideline Three has been met. In particular, I disagree with his interpretation of the Hoaglin and Glickman study, which he says supports the homogeneity assumption. As noted above, it can be used to support a conclusion that variance is a serious problem with the synthetic estimation model.

I agree that Guideline Six can be met.

I disagree that small differences between alternate sets of adjusted figures are no bar to adjustment, given the requirements to adjust to the block level with distributive accuracy.

Recommendation of Kenneth W. Wachter

Summary of the Recommendation

Wachter recommends against an adjustment. He "conclude[s] that the requirements for accuracy, state and local usability, and robustness articulated in Guidelines One, Two, and Three are not met by the adjusted counts. The broader considerations in Guidelines Four through Eight also, on balance, do not favor a decision to adjust. [He] therefore recommend[s] against adjustment of the 1990 U.S. Census counts." [63]

On Guideline One, Wachter concludes that the adjusted counts are not satisfactory. Although:

> evidence indicates that the adjusted counts are more accurate at the national level, the relative sizes given by adjusted counts are probably less accurate for a number of [S]tates and surely less accurate for a substantial fraction, possibly a majority, of local areas for which [c]ensus counts are to be used." [64]

As a preface to detailed sections on Guideline One, Wachter makes several pages of general observations:

The adjustment of a census is difficult as it is a matter of changing the counts for 6.8 million blocks. A post-enumeration survey (PES)-like survey is usually used to generalize up from sample totals to population totals; for such a use the absolute size of the sample rather than the fraction surveyed would limit the accuracy that could be achieved. The PES is used by the census to generalize down, which is a much more demanding process.

Three things must happen for the PES to be successful. The PES operation must be good, the people missed in the Census have to be reached by the PES, and the reasons why people are missed must be knowable so that one can extrapolate from the people and places analyzed to all the rest, for the PES to improve the census enumeration. The first has happened, the second has not, and the third is in doubt. [65]

The quality of the PES is high. There are problems and limitations but no disasters. Thus the first criterion is met.

A substantial portion of persons missed, net, by the census were not within reach of the PES. Discrepancies between estimates of national undercount between the PES and demographic analysis by age and sex for blacks and non-blacks cannot be explained away by plausible allowances for uncertainty. Half the black males who are missed, net, in the census are being missed, net, in the PES. There is no direct information on the distribution of these people from place to place.

As to the third criterion, the answers are not yet clear-cut. There is insufficient homogeneity at different levels of disaggregation for post stratum for the adjusted numbers to be usable. Erroneous enumerations are numerous and prominent in the adjustment picture. Block level data and district office data do not support the assumption of homogeneity.

Different smoothing procedures should lead to similar answers with respect to adjusted versus enumeration counts, but they do not: they lead to markedly different answers. [66]

Combining census and PES data produces results that are better than either alone only if we know enough about the precision and accuracy of each part. This is an empirical, not an *a priori*, question.

His personal experience with census enumerators and PES enumerators suggests that, contrary to common wisdom, census enumerators may very well have done a better job than PES enumerators in a significant class and number of cases.

> *[He] do[es] not believe that any highly aggregated index or loss function is appropriate for summing up overall accuracy.* It is informative to understand how much the outcomes of calculations with different versions of such aggregated indices differ. But the choice among them is not a scientific choice. Each such index involves implicit value judgments about different sorts of error. For example, each index determines whether a few large errors are more serious than a great many smaller errors. *Whether we agree with a particular tradeoff is a matter of personal and political values. It should not be disguised as science.* [67]

[60] ukey, page 2.

[61] McGehee, page 6; Tarrance, pages 4–6.

[62] See the discussion in Guideline One above.

[63] Wachter, page 2 of cover letter.

[64] Wachter, page 1, emphasis in the original.

[65] Wachter, pages 1–2.

[66] Wachter, page 3.

[67] Wachter, page 3, emphasis in the original.

The census is the source of small-area data, so accuracy at that level has a special claim although some sensible balance of concern and perspective for level of detail is required.

In the first section devoted to Guideline One, Wachter considers national discrepancies between the PES and demographic analysis. There is a national undercount, although Wachter takes issue with the uncertainty intervals about the point estimates of the undercount. There is also credible evidence of a differential undercount. Although the evidence from the demographic analysis and the PES agree as to the existence of broad differentials, "the evidence as to the pattern by age and sex for blacks and non-blacks does not agree." [88] According to the demographic analysis, a high undercount rate for black adult males, aged 20–64 exists. This does not occur in the PES which means that *"a large portion of the people probably missed by the Census were also missed by the extrapolation from the PES that produced the adjusted counts."* He calls these people "unreachable." [89]

Wachter estimates the numbers of unreachable people to be large, perhaps half-a-million. Since nothing is known about their location, the huge numbers of "unreachable" people mean relative population sizes based on adjusted counts cannot be shown to be more accurate than those based on census counts at any subnational level.[90]

Wachter then turns to patterns in the estimates of net undercounts for post-strata. The patterns of adjustment factors for the 1392 post-stratum groups show regular patterns at higher levels of aggregation, but unexpected complexity when examined stratum by stratum, suggesting heterogeneity where there should be homogeneity. Analysis "for aggregates mask a large amount of diversity within groups, and the story of census coverage, at a level of fine detail, is more complicated than one would hope." [91]

Wachter then turns to the proximate determinants of net undercount. He finds that "erroneous enumerations account for a large portion of the variations in net undercounts across areas and post-strata." [92] Erroneous enumerations play a powerful role in determining the net adjustments to the counts, and this role is masked by smoothing adjustment factors which is probably unjustifiable.[93]

Wachter suggests that variation in erroneous enumeration could be the result of coverage improvement programs. The evidence that can be gleaned from comparing the cities of Detroit and Chicago is mixed. The main conclusion that can be drawn is that "erroneous enumerations are extremely varied. . . . [However,] lumping Detroit and Chicago together in the same post-strata, as the PES does, ignores sizable differences in coverage patterns." [94]

Wachter says that strong correlation between erroneous enumerations and omissions is insufficiently understood, even though it contributes substantially to the size of net undercounts. Since erroneous enumerations exceed omissions in a good number of post-strata, there will be a goodly number of downward adjustment. Thus *"people who themselves filled out their Census forms correctly may be 'minused out' of the Census to compensate for others who were erroneously enumerated"* to calculate an adjustment. "There may be no statistical objection to such a process. But on a human level it is offensive." [95]

Wachter asserts that there remain uncertainties in the demographic analysis, although it has been much improved.

Wachter states that the total error model does not mean all relevant errors for assessing the accuracy of a PES are included. Rather it addresses errors at the level of the evaluation strata only and, furthermore, treats them separately with no joint error structure. There is no simple way to generalize from the evaluation strata to small areas.

The approach is novel, pioneering and controversial. Thus, the "confidence intervals" around error components are not what statisticians usually mean by confidence intervals. The total error model actually estimates only a portion of the possible sources of error in undercount estimates. Components missed are of unknown magnitude. Stratification is applied inconsistently and some of the uncertainty estimates are themselves subject to large uncertainty. The total error model is too optimistic with respect to uncertainties attributed to imputation.

For Wachter, the main lessons drawn from the total error model are that the confidence intervals for most of the non-minority strata are compatible with zero net undercount, but the intervals for all the minority evaluation strata are not. The higher estimated undercounts are subject to high estimated biases.[96]

Several critical aspects of the total error model results are then discussed by Wachter, beginning with correlation bias or "catchability error." The correlation bias assumptions used are not realistic when applied to the PES. People stay out of the Census and the PES not by chance, but because they want to. Dual system estimation depends on chance mechanisms. There are many ways to allocate the twice-missed people. Whether the choice made is good is entirely speculative. How the measurement of variance in the total error model reflects correlation bias is not clear. It is better not to attempt any formal allocation of unreached people to local jurisdictions because of these problems.

Wachter next turns to matching and imputation studies. These studies of matching error give estimates of false non-matches that are too low by the very nature of their design. A small test on step-children illustrates the point that because matchers simply apply rules, they may miss true matches.[97] The effects of imputation may also be larger than the evaluation studies indicate. Wachter uses a sensitivity analysis to indicate the bounds on the effects of imputation. It shows that a great deal rests on the correctness of the assumptions in the imputation, but since these assumptions have not been examined, the measures of variance are too low.

On Guideline Two, Wachter sees "substantial obstacles to using adjusted data for Congressional reapportionment" and concludes that adjustment procedures are not well suited for coping with local heterogeneity in census undercounts. Firm conclusions cannot yet be drawn as to the extent of local heterogeneity and its implications for the accuracy of adjusted local counts.

Wachter shows by example that depending on how imputation is done seats could shift between States in a variety of ways. In estimating adjusted state population counts, adjustment factors based only on within-State data, rather than factors including across state data affect the distribution of Congressional seats as well. Among the five methods tried by Wachter, each apportionment was different and eleven states either gained or lost a seat relative to the census in at least one of the methods.

[88] Wachter, page 8.

[89] Wachter, page 7, emphasis in the original.

[90] Wachter, page 9.

[91] Wachter, page 10.

[92] Wachter, page 11.

[93] Wachter, page 10.

[94] Wachter, page 11.

[95] Wachter, page 14, emphasis in the original.

[96] Wachter, page 17.

[97] Wachter, page 21.

Wachter points out that there is acknowledged lack of homogeneity w[illegible] post-strata. The issue is whether t[illegible] evere to make adjustment lo[illegible] infeasible. Very little is known about local heterogeneity. Experiments at the block level give ambiguous results with respect to the balance between improvements and worsenings of counts when adjustments are carried down to the block level. However, Wachter concludes that local heterogeneity is a serious problem for adjusting the counts at district office levels and that perhaps a majority of units could be made worse by an adjustment.

Wachter's experiments and analysis convince him that studies of local-level adjustment have "scarcely begun to scratch the surface" of the issue of how local heterogeneity has an impact on adjustment.[98] His block level analysis leads to more puzzles than answers.

On Guideline Three, Wachter finds that reasonable alternatives to one aspect of the smoothing model lead to significantly different adjustment factors and thus the adjustment factors cannot be considered robust. He finds that smoothing has been the most problematic part of the PES and that the smoothing has had more of an effect on the final adjustment than can be easily justified. The effect of deciding to use pre[illegible]thed rather than unsmoothed va[illegible] in computing smoothed adj[illegible]nt factors is to raise many adjustment factors by several percentage points, some by more than six percentage points. The changes introduced into the adjustment factors are of the same order of magnitude as the sizes of the adjustment factors themselves.[99] Decisions about pre-smoothing make a large difference and so alternate methods leading to different outcomes seem equally reasonable. In fact, pre-smoothing seems to run the risk of "loading the dice." [100]

Wachter argues that pre-smoothing of variances changes variances in ways that are counter to what one ought to do: reducing large variances increases the weight assigned to empirically unstable factors; increasing small variances reduces the weight assigned to stable factors. In addition, the variance smoothing process is not directed at making covariances more accurate. Furthermore, the choice among regression models is arbitrary in the sense that there is no reason to choose among them, yet the results each set produces differ from one another substantially. Finally, smoothing affects not only adjustment factors, but higher level aggregations of data.

Wachter observes that the effects of the selection of variables for the regression part of smoothing are not negligible but they are not a central issue.

On Guideline Four, Wachter feels that an adjustment would reduce the stake that individuals, civic leaders and Congressional representatives would have in coverage improvement efforts. Adjustment would increase the political leverage of technical decisions and extra efforts to guarantee the Census Bureau's independence and objectivity would be required.

Wachter offers no guidance on Guideline Five.

On Guideline Six, Wachter states that sufficient data are available for a reasoned decision on adjustment.

On Guideline Seven, Wachter says that disruption is likely as a result of an adjustment, but this should not be decisive for the adjustment decision.

On Guideline Eight, Wachter sees no difficulty in meeting professional standards of the scientific community. The details of the adjustment decision tell against its understandability by the general public. Some dismay when an understanding of what adjustment really is should be anticipated, if the decision is to adjust.[101] Adjustment will have victims.[102]

Evaluation of Recommendation

I agree with Wachter's point that the PES, even if it yields results more accurate at the National level, doesn't improve the distribution of population over the results of the census enumeration totals due, in part, to "unreachable" people; among other factors.

I agree with the argument that a good PES is not a sufficient reason to adjust the census. I agree that Wachter's two other conditions are not met, *viz*, people who were missed must be reached, and why they are missed must be knowable.

I agree that Wachter's elaboration of the problem of correlation bias provides insight into why the adjusted counts produced from the PES may be distorted by correlation bias, and not simply underestimate the undercount. There are simply people who are unreachable, and determining why they are unreachable is an insoluble problem.

I agree with the analysis of discrepancies between the PES and demographic analysis.

I agree that the total error model does not include all, or necessarily even most, sources of error. I agree with the criticism that the confidence errors around the components of the model are speculative, and not uncontroversial among statisticians. Pointing out that higher estimated undercounts are subject to higher estimated biases casts serious doubt on the quality of these PES estimates.

I agree when Wachter states that the total error model does not mean all relevant errors for assessing the accuracy of a PES are included. I agree with him as he goes on to say, "Rather it addresses errors at the level of the evaluation strata only and, furthermore, treats them separately with no joint error structure. There is no simple way to generalize from the evaluation strata to small areas. The approach is novel, pioneering and controversial. Thus, the 'confidence intervals' around error components are not what statisticians usually mean by confidence intervals. The total error model actually estimates only a portion of the possible sources of error in undercount estimates. Components missed are of unknown magnitude. Stratification is applied inconsistently and some of the uncertainty estimates are themselves subject to large uncertainty. The total error model is too optimistic with respect to uncertainties attributed to imputation."

I agree with the discussion of Guideline Two that more work is needed to determine the homogeneity problem at the local level.

I agree with Wachter's conclusions with respect to robustness that interpret findings concerning the output from different models as raising questions about robustness at lower levels of disaggregation. In addition, smoothing is correctly identified as a significant factor affecting outcomes for higher level aggregations of data.

Recommendation of Kirk M. Wolter

Summary of Recommendation

Wolter recommends an adjustment. His analysis relies on the joint paper co-authored by Ericksen, Estrada, Tukey, and Wolter. The corrected counts, as required by Guideline One for an adjustment, are more accurate in both level and distribution at the national, state, and local levels.

Wolter finds Guideline One to be the pre-eminent guideline. His conclusion that the corrected counts are more accurate is based first on the observation that the post-enumeration survey (PES) is superior to the census by

[98] Wachter, page 30.

[99] [illegible] page 36.

[100] [illegible] page 37.

[101] Wachter, page 40.

[102] Wachter, page 40.

virtue of the design of matching operations and interviewer training and second, because a survey can be more tightly controlled than a census. The evaluation studies demonstrate that missing data, quality of Census day addresses, fabrication, matching, erroneous enumeration measurement, and balancing sources of error were controlled in the PES to very low levels. Correlation bias, while not so well controlled, is an error such that the PES estimates are still closer to the truth. Random error does not affect the utility of PES estimates.[103]

Wolter's rationale for preferring the adjusted counts includes four major points: (1) PES estimated undercounts agree with expectations and with demographic analysis; (2) the total error analysis demonstrates that corrected counts are more accurate for states, counties, and other similar areas; (3) corrected counts for evaluation strata, which are relatively homogeneous, offer even more improvement than they did for states, especially in comparing five minority with eight non-minority strata and central city versus non central city strata; and (4) if the stratum-level undercount rates are accurate, then the corrected counts for local areas must be an improvement on uncorrected counts.[104] This latter result is based on the Wolter/Causey paper that is appended to the coauthored report as Appendix G. Wolter also cites the plaintiffs co-authored report.

On Guideline Two, Wolter states that the bureau is capable of producing adjusted counts down to the block level, so the first part of the Guideline is satisfied. As to accuracy at small area levels, Wolter notes that, synthetic estimates of the kind used on the 1990 census can improve accuracy at small area levels so long as measured undercounts at aggregate levels tend to have smaller error than the original enumeration at aggregate levels. In support of his position, he again cites the Wolter/Causey paper. The Bureau's P12 study also offers evidence that the adjusted counts are superior to the census counts at the local level.

On Guideline Three, Wolter argues that the PES adjustment procedures were sufficiently prespecified to satisfy the guideline. The three instances where the procedures were not prespecified were "treated with a high degree of objectivity and professionalism."[105] The Hoaglin and Glickman report demonstrates that corrected counts are robust to variations in reasonable alternatives i.e. the smoothing component of the overall PES process. The Census Bureau P1 study demonstrates that the PES undercount estimates are insensitive to differences in the manner of handling missing data.

On Guideline Four, Wolter states that "it is virtually impossible to say anything about the public's cooperation with the 2000 census."[106] The National Opinion Research Center (NORC) study indicates that the average American doesn't understand adjustment, plans to participate in future censuses, and that the adjustment decision, one way or the other, would have little effect. Other countries have instituted adjustment into their censuses with no adverse effect on public participation. Using the most accurate counts is the best way to handle the perception that the adjustment decision is a politically motivated act because Wolter believes that no matter what the decision is—it will be perceived as politically motivated.[107]

On Guideline Five, Wolter acknowledges that he is not a lawyer, but his understanding is that there is no legal ruling that stands in the way of an adjustment.

On Guideline Six, Wolter finds that the necessary data upon which to base the adjustment decision are sufficient, complete and available, and provide a sufficient basis for the adjustment decision.

On Guideline Seven, Wolter finds that the States have been alerted to the possibility of adjusted counts, and can deal with it. The Census Bureau analyses of misapportionment suggests that the original enumeration would misapportion seats more than the adjusted counts. Thus, not adjusting could be viewed as generating more disruption. Wolter is "unaware of any aspect of the 1990 correction process that would cause a truly calamitous disruption of the political process."[108] No part of the correction process has been arbitrary because scientific principles have guided the effort.

On Guideline Eight, in Wolter's view, there is a clear rationale for certifying the correct counts and the Bureau's documentation of the process has been satisfactory. The Bureau and the Department should be able to articulate clearly the basis for the adjustment decision

Evaluation of Recommendation

I do not agree that the PES counts are superior to the census counts. The four points of Wolter's rationale for believing the PES superior are flawed. Contrary to Wolter, PES undercounts do not agree with expectations, or the demographic analysis.[109] For example, the PES misses half a million black males which demographic analysis says are in the population. The total error analysis deals with numeric, not distributive accuracy. Thus, whatever it concludes about accuracy is not to the point of the form of accuracy which must be demonstrated.[110] The homogeneity assumption is in doubt.[111] There is not agreement on the inevitability of increased accuracy at lower levels, notwithstanding a certain degree of accuracy at broader levels.[112]

I do not agree that the synthetic estimate evidence in support of Guideline Two is clearcut, as Wolter states. In particular, P12 casts serious doubt on the homogeneity assumption necessary to a successful synthetic adjustment.[113]

I do not agree with Wolter's interpretation of the evidence with respect to robustness. I believe that the Hoaglin and Glickman report demonstrated that thirteen different models give thirteen different answers. An outcome of that kind is not robustness in the practical sense demanded by this guideline.

I agree that Guidelines Four and Five are no bars to an adjustment decision. On Guideline Six, I note that some panelists feel there is concern that census studies were not sufficiently analyzed in the time frame agreed to in the stipulation and order.

I do not agree that the Census Bureau analyses of misapportionment of Congressional seats are adequate.[114] I do not agree that there is clear consensus that the states can deal with adjusted counts.[115] In my view, while this does not bar adjustment, it remains a consideration to be reckoned with.

[103] Wolter, page 4.

[104] Wolter, pages 4–8.

[105] Wolter, page 9.

[106] Wolter, page 11.

[107] Wolter, page 11.

[108] Wolter, page 15.

[109] See the discussion in Guideline One above.

[110] See the discussion in guideline 1 above.

[111] See Appendix 2.

[112] Wachter, pages 2–3.

[113] See the discussion of distributive accuracy in Guideline One above.

[114] See the discussion in Guideline One above.

[115] See appendix 12.

Recommendation submitted jointly by Eugene P. Ericksen, Leobardo F. Estrada, John W. Tukey and Kirk M. Wolter

Summary of the Report on the 1990 Decennial Census and the Post-Enumeration Survey

The authors begin by considering the enumeration. The census differentially undercounts Blacks, Hispanics, Asians, and Native Americans. The Black undercount has been documented since 1940; the Hispanic since 1980. Differential undercounting is a result of the way the census is taken because it works best for "middle-class suburban" households and worst where living conditions are different. Undercount is strongly negatively correlated with the mailback rate.[116]

The authors state that the original enumeration of the population in 1990 experienced a staggering array of problems. The mail response rate was low, coverage differed between minorities and non-minorities, enumerators gathered less accurate information in cities than in other areas, and nonresponse follow-up operations had a high proportion of last resort and non-data defined responses. The difficulties in enumerating urban areas can be seen from the data. In large city offices 20% of all nonresponse followup was last resort or closeout versus 12% in small city/suburban offices and 11% in rural offices.[117]

The authors claim that last resort and closeout information is incomplete and often inaccurate. More than one-third of all last resort information and 44% of all closeout cases were estimated to be erroneous.[118] Re-enumeration of households originally enumerated by last resort or closeout showed serious errors in certain problem offices. In a national survey of 1,000 one-person households there was between a 20% and 25% error rate depending on the measure used.[119]

The authors say that coverage improvement programs, while adding people to the count, were frequently in error. For example, more than 630,000 of the 2.1 million persons added through vacant/delete either should not have been added at all or should have been added at a different place. More than half (53%) of the persons added to the count through the parolee/probationer check were estimated to have been added in error. Overall, the coverage improvement programs failed to do what they were supposed to—accurately add a substantial number of persons to the census count and the differential undercount remained after the programs had been completed.[120]

In addition to adding error to the count, the authors argue that the coverage improvement programs failed to find the estimated 19.2 million persons actually missed by the census. The "Were you counted" campaign and the Housing Coverage Check and Local review added only 200,000 and 300,000 persons, respectively, to the count. The low number of accurate additions left intact and possibly increased the differential omission rates by race and type of area that had already existed.[121]

The authors next turn to demographic analysis. Demonstrating through demographic analysis that a black non-black differential undercount exists for every census since 1940, the authors conclude that a black non-black differential undercount exists by virtue of demographic analysis in the 1990 decennial census.[122]

Next, the authors turn to the post-enumeration survey (PES). The PES is the mechanism designed by the Census to determine the extent of, and correction for, census error. The post-enumeration survey has demonstrated the differential undercount of the minority population and solved the major error of the original enumeration, which was the inappropriate shifting of shares of population from areas with many minorities to areas with fewer.

The authors state that the PES was a high quality survey. Completed interviews were obtained 99% of the time for the total PES sample, and for major geographic and racial subgroups. Proxy interviews accounted for 2.4% of the total sample, with little variation in this rate across subgroups. Only 1.5% of the P-sample were unresolved in the matching operation, and only 0.9% of the E-sample. There was little subgroup variation.

The authors use three criteria to evaluate the success of the PES: consistency with expectations of the distribution of the undercount (*i.e.* rates of omission and erroneous enumeration should be higher where census taking was more difficult) and the results of demographic analysis; the P studies (looking at missing data and the outcomes of rematch studies especially); and the possible shifting of population if net undercount rates were altered as a result of the P studies.

The authors state that PES results were consistent with substantive expectations especially when compared with demographic analysis.[123]

The authors' examination of P studies focused on four problems: The effect of variation in assumptions on how to treat missing data; problems due to matching error; problems with census day address misreporting and matching error for movers; and correlation bias. Assumptions about how to treat missing data had little effect. Because the numbers of movers were small, mover matching error had little effect. Correlation bias was a major source of error. Its effect tends to be to reduce estimated undercount. Evidence from evaluation poststrata research shows that adjustment increased the minority share of the nation's population by 0.8%, from 21.4% to 22.2%. The total error model showed a shift of 0.76%.[124]

The next major area considered by the authors was the smoothing of the adjustment factors. They consulted with David Hoaglin to evaluate the impact of the decisions on carrier variable choice, how to smooth variances and covariances of raw adjustment factors before calculating the regression, and how to weight individual observations when calculating the regression.

Hoaglin identified how to smooth the variances before using them to weight observations in the regression calculations and how to smooth the covariances before using them for the same purpose as key decisions.

Hoaglin fitted thirteen different regressions. The first nine were based on three strategies for smoothing variances and three strategies for smoothing covariances ($3\times3=9$); a tenth alternative was suggested by a Panel member; finally for comparison purposes he considered equal weighting of observations; weighting according to raw variances and covariances; and weighting according to raw variances, replacing the covariances by zero.[125]

After considering various alternative "stopping rules" for the "best subsets regression," Hoaglin chose a "back-2" stopping rule which uses apparently the best subset among those involving two fewer carrier variables than are in the set that minimizes the ratio residual mean square/residual degrees of freedom.

Hoaglin used two strategies to test whether the decisions had serious impact on the estimates: The first strategy used the difference in fitted values from each pair among the 13 choices and differences between the 13

[116] Ericksen, *et al.*, pages 1–2.

[117] Ericksen, *et al.*, pages 4–5.

[118] Ericksen, *et al.*, page 6

[119] Ericksen, *et al.*, page 6.

[120] Ericksen, *et al.*, pages 7–8.

[121] Ericksen, *et al.*, page 8.

[122] Ericksen, *et al.*, pages 10–11.

[123] Ericksen, et al., pages 13–14.

[124] Ericksen, et al., pages 12–16.

[125] Ericksen, *et al.*, page 18.

and the Bureau's regression fit; while the second strategy used the reallocation of population shares among the 13 evaluative post strata.

Hoaglin stated that alternative smoothing models produced estimated population share gains for minorities that closely "surround the Bureau fit," ranging from 0.48% to 0.77%[126]

Next the authors considered errors for large and small areas. In looking at the differences in errors for large and small areas, they concluded that the total combined error increases as the size of the group decreases (e.g., the combined errors for 5 million blocks will be larger than the combined errors for 1,392 poststrata), and consequently the improvement in amount due to adjustment would be nearly the same for larger and smaller groups—the improvement in percentage terms decreases, but does not change sign, as the groups become smaller.

The authors stated that since the expected CV for a sampling stratum is 1.4%, they were more likely to expect improvements for those areas where undercounts are especially high or especially low. It is these extreme cases where most of the benefit of adjustment is to be expected. Improvements in quite large areas thus prophesies improvements in very small areas, as well as in intermediate areas.

The authors' major conclusions are that error in the uncorrected census was very high; this error disproportionately affected Blacks, Hispanics, Asians and Native Americans; and the PES derived data can be used to correct the census and substantially reduce the differential undercount and improve accuracy at both national and local levels.

Evaluation of the Report on the 1990 Decennial Census and the Post-Enumeration Survey

I do not find the discussion of the quality of the census relevant. Guideline One stipulates that the census is the standard. Thus, irrespective of the flaws in the census, Guideline One precludes adjustment unless the adjustment is shown to be better than the census by convincing evidence.

I do not agree with the statements in discussions of the PES claiming that PES results were consistent with expectations when compared to demographic analysis is made. There were sizable, and unexpected differences between the PES and demographic analysis which indicate that a PES based adjustment would be inadequate. [127]

I do not agree with the interpretation of the Hoaglin materials. The authors' interpretation misses the point. The issue is not whether the thirteen different outcomes fluctuated around a Bureau estimate of "truth" derived from the PES and are thereby defined as demonstrating sufficient robustness. The very fact of such a variety of outcomes is precisely the lack of robustness that is of concern when using a model based synthetic adjustment at a low level of geography.

The authors state that the expected CV for a sampling stratum was 1.4%. The expected CV was .7%.

I do not agree that PES derived data can be used to correct the census and substantially reduce the differential undercount and improve accuracy at both national and state levels. [128]

SECTION 4—DECENNIAL CENSUS PROCEDURES

In this section I provide documentation for the procedures used to conduct the decennial census, the post-enumeration survey, the evaluation of the post-enumeration survey, and the evaluation of the demographic analysis. Additional information on the post-enumeration survey evaluation program and demographic analysis will be found in appendix 3.

1990 Census of Population and Housing: The Bicentennial Census of the United States

Planning for the 1990 Census began in 1984, with planning activities, testing, and preparatory operations occupying the remainder of the decade. Data were collected in 1990, and, as required by law, State population and apportionment totals were delivered to the President on December 26, 1990. The total population count transmitted to the President was 249,632,692, composed of a resident population of 248,709,873 and an overseas population of 922,819.

The Census Bureau was also required by law to deliver redistricting counts and maps to State redistricting officials no later than April 1, 1991. This was done. While the Census Bureau met its two legal mandates for the delivery of apportionment and redistricting data—two of the most important uses of census data—the 1990 census is not considered completed until all planned census data products have been released. Final products will be released in 1993.

The 1990 census involved enumerating 249,632,692 people in more than 100 million housing units, and collecting a full range of characteristics about each person. Extensive planning and preparation, the successful recruitment and employment of hundreds of thousands of temporary census workers, and an automated management information system to keep track of operations were required to complete the census on time and within budget.

Planning and Preparation

The Census Bureau designed the 1990 census keeping in mind the special problems that arise in the census-taking process, as well as constraints of time, budget, and the need to protect individual confidentiality. Plans incorporated the lessons learned from previous censuses. The plans were tailored to implementation and management by a temporary work force in a compressed time frame. Extensive testing was conducted so that hard evidence could be gathered on the utility of new procedures and techniques. The testing also allowed new procedures and techniques to be refined and adjusted.

Formal planning for the 1990 census began in FY 1984. This early start allowed the Bureau to begin major testing of proposed design features earlier for the 1990 census than for the 1980 census (1984 *vs* 1976), and to conduct more major tests of proposed features than for prior censuses (e.g., 7 for 1990 *vs* 5 for 1980). Improvements were made in every phase of census-taking. Some were aimed directly at overcoming operational, control, and timeliness problems identified in 1980 census operations. Others were intended to increase the cooperation of hard-to-enumerate groups. These improvements are described in detail in "Planned Improvements in the Counts for the 1990 Census," April 1989, Bureau of the Census. Improvements included:

- An expanded promotion campaign aimed at hard-to-enumerate groups. For example, for the first time, the Bureau used minority advertising campaigns designed by minority firms, in addition to a more traditional general-audience campaign.
- More cooperation between the Census Bureau and state and local governments. For example, the Census Bureau improved and expanded the Local Review Program, which gives local officials an opportunity to review census counts, by providing training on how to participate in the program, and by instituting two phases of review instead of one, as was the case for the 1980 census.
- Efforts intended to make it easier for people to respond ... census questionnaires. For example, the Bureau expanded questionnaire assistance operations for 1990 by offering toll-free

[126] Ericksen, et al., pages 17–19.

[127] See the discussion in Guideline One above.

[128] See the discussion in Guideline One above, where the deficiencies in distributive accuracy of an adjusted count, using Census Bureau procedures, are detailed.

telephone assistance in English, in Spanish, and in six Asian languages, and by sending out multilingual "early [illegible]" flyers about the census in selected

[illegible]ailoring census procedures to deal with special or unusual situations. For example, enumerators delivered questionnaires to public housing developments, and the Bureau hired public housing residents to deliver the questionnaires and conduct outreach activities at the same time.

• A greatly increased amount of automation in the census. For example, an automated management information system, in conjunction with an automated address control file, enabled home office control and monitoring of the 1990 census to deal with developing problems early and rapidly.

• Implementing an automated geographic control system—called TIGER—in cooperation with the U.S. Geological Survey. The TIGER System solved one of the most serious problems of the 1980 census—late, inconsistent, and illegible maps. The TIGER System assured accurate and timely maps and geographic files for the 1990 census.

The 1988 dress rehearsal was the capstone of planning efforts; it was preceded by 5 years of consultation with data users and formal tests of alternative procedures and questionnaire content of the kind just [illegible]bed. The Bureau consulted with a [illegible]ange of data users, including [illegible]ity organizations, planners and academics, business leaders, representatives of private organizations, state and local officials, and Federal agencies.

Once the basic plan for the census, including improvements, was determined, the Census Bureau began to prepare for 1990 data collection and processing. These preparations included map-making, questionnaire printing, address list construction, setting up a field structure of over 500 offices for data collection and processing, procuring and installing automated equipment, and preparing promotion materials.

A critical activity was preparation of a precensus address list. This list was used to determine which housing units had or had not returned a questionnaire in areas where householders were instructed to return their questionnaires by mail. In all, some 100 million addresses were compiled before the census from purchased lists, field canvassing by census enumerators, and a series of overlapping checks and update operations by census workers, the U.S. Postal Service, and review by [illegible]fficials.

By March 1990, all preparatory activities had been completed and the data collection phase of the census, which involved attempting to get a completed questionnaire for every person and housing unit in the Nation, was set to begin. (Enumeration of remote areas of Alaska had begun a few weeks earlier in order to complete the enumeration before the Spring thaw.)

Basic Enumeration Procedures

The 1990 census was planned to be a multiphase and incremental process that was to determine the population as of April 1, 1990. Except for remote areas of Alaska, questionnaire delivery or mail-out occurred in March 1990, but the enumeration was not intended to be over then. The Census Bureau built into the census process programs to follow up on housing units that did not return a questionnaire and to ensure that every reasonable effort was made to enumerate every housing unit. These programs extended well after April, into the fall of 1990.

90 percent of the housing units were expected to complete questionnaires and return them by mail. Two procedures were used in such mail-back areas—mail-out/mail-back and update/leave.

For the remaining housing units, householders were instructed to hold their completed questionnaires for enumerator pick-up. This procedure was called list-enumerate. Other special procedures were designed to enumerate persons who lived in group quarters (such as college dormitories and military barracks) and persons who had no usual residence.

Mail-Back Areas

Mail-Out/Mail-Back

The mail-out/mail-back procedure was used for large cities, suburban areas, and some smaller cities, towns, and rural areas where mailing addresses were house number and street name. In all, about 83 percent of U.S. housing units were in mail-out/mail-back areas. Mail carriers in these areas delivered addressed questionnaires on March 23, 1990, and householders were asked to mail back completed questionnaires by April 1, 1990. Five out of six housing units received a short form containing *only* the questions asked of all housing units; one out of six housing units received a long form with additional questions. One week after mail-out, a post card was sent to each housing unit reminding persons to fill out the questionnaire and return it as soon as possible. This was in addition to the multiple-component promotion campaign, then at its peak.

The USPS returned some questionnaires to the Census Bureau as "undeliverable." The Bureau added a special operation to have census enumerators deliver by hand as many of the "undeliverables" as possible. The remaining housing units did not receive a mailing piece at this time, so they were enumerated during nonresponse follow-up (see below).

Update/Leave

The update/leave method was used in rural areas in the South, Midwest, and Appalachia, where mailing addresses are rural-route designations, or where many householders pick up their mail at lock-boxes. These areas contain about 11 percent of the housing units in the Nation. Here, census enumerators, rather than the USPS, delivered the census questionnaires and, at the same time, updated the address list. This operation began in early March 1990 and continued throughout that month. Just as in mail-out/mail-back areas, householders in update/leave areas were to complete and mail back their questionnaires by April 1, 1990. Again, most units received a short form, but a small pre-designated sample received the long form. Householders in these areas also received a reminder postcard asking them to return their questionnaires.

List/Enumerate

The list/enumerate, or door-to-door method, was used for about 6 percent of the Nation's housing units. These units were primarily in very remote and sparsely settled areas. There was no precensus address list for these areas. Mail carriers delivered unaddressed short-form questionnaires on March 23 and, beginning about April 1, census enumerators went door-to-door listing addresses, picking up completed questionnaires or filling out questionnaires as necessary, and administering the long form at a sample of these units.

Special Procedures

Special place enumeration took place in March and April, 1990. Special places include group quarters, such as boarding houses, nursing homes, dormitories, rectories, convents, hospitals, etc. Enumerators visited these places to collect information from each resident. About 2 weeks before Census Day, the Census Bureau also conducted a Street and Shelter enumeration (S-night) to collect information from components of the homeless population. The first phase of this operation focused on enumerating persons staying in shelters for the homeless, while the second phase focused on enumerating homeless persons living outside of shelters, for example, on the street.

There were two additional components of special place enumeration: Transient enumeration and military enumeration.

• During transient enumeration, census workers visited travel places where guests are unlikely to have been reported at their usual place of residence, or where guests are unlikely to have a permanent residence. These places include YMCA's, YWCA's, youth hostels, commercial campgrounds, etc.

• For military enumeration, special procedures were used to count domestic military and maritime personnel. Military bases and vessels were self-enumerating. In these instances, bases appointed a senior commissioned officer to serve as the enumeration project officer.

Questionnaire Receipt

Some households received a short questionnaire containing only the questions asked of all households, while others received a long form containing additional questions. About 17 percent (or a sampling rate of about 1-in-6) of the households received the long form. However, in places with an estimated 1988 population of less than 2,500, the sampling rate was 1-in-2. Based also on precensus estimates, very populous census blocks had a sampling rate of 1-in-8. All other areas had a sampling rate of 1-in-6.

Once questionnaires had been delivered, forms began to arrive by mail in district or processing offices serving each area. Mail returns for some areas went to a processing office for check-in. For most areas, mail returns, as well as questionnaires completed by enumerators during list/enumeration or special place enumeration, went directly to a district office. Both processing offices and district offices used automated equipment to check in forms by bar code scanning of the return envelope. The associated address in the automated address control file was then coded to show that a questionnaire had been received for that unit. At the conclusion of the check-in phase, each listing not coded represented a case that would have to be visited by an enumerator during nonresponse follow-up.

Nonresponse Follow-up

The Census Bureau followed up every housing unit for which a questionnaire was not returned. Daily reports on the mail return check-in rates for each district office were transmitted to headquarters through the automated management information system. This information was used to project the likely workloads for nonresponse follow-up. This overall workload was expected to require over 250,000 temporary enumerators to visit 30 million units over a 2 month-period. By the end of April, the Census Bureau had to estimate the number of persons it needed to hire, and to begin preparing lists of addresses that had not returned a questionnaire. The mail response rate was 63 percent, lower than the projected 70 percent. As a result of this, the Census Bureau hired more enumerators than it had originally planned for nonresponse follow-up.

The Census Bureau completed nonresponse follow-up for the 1990 census substantially earlier than had been the case for the 1980 census, despite a larger workload. Recruitment goals were met despite the need for more workers engendered by the low mail response rate, and in spite of lower levels of general workforce unemployment than had been the case for the 1980 census.

During nonresponse follow-up, enumerators were required to make up to six attempts to contact a household member and complete a census questionnaire. If this was not possible after three personal visits and three telephone calls at different times and on different days, the enumerator attempted to obtain at least basic information on household member(s) from knowledgeable sources, such as neighbors or building managers.

Because the nonresponse follow-up had to be completed quickly so that other operations could be conducted, each district office was authorized to begin a final phase of nonresponse follow-up once 95 percent or so of the operation had been completed. During this phase, enumerators made one more visit to each remaining case to obtain as complete an interview as possible.

Coverage Improvement Efforts

Basic data collection activities included various steps designed to improve census coverage. Among these were special promotion and outreach efforts, better address listing procedures, extra efforts to increase mail returns, follow-up on all housing units that did not return a questionnaire, better management of and pay for enumerators, etc. But after basic data collection, census plans also included additional special programs to improve the population count that went beyond standard procedures.

These additional coverage improvement programs, which represent the Census Bureau's policy of giving everyone several opportunities to be included in the census counts, added about 5.4 million persons to the census counts, or about 2.2 percent of the total enumerated population.

Such coverage improvement programs included: (1) The 100-percent recheck of vacant housing units or those identified as uninhabitable or nonexistent; (2) the "Were You Counted?" campaign, an opportunity for people who thought they might have been missed to call in or fill out a census form printed in the newspaper; (3) the parolee and probationer check, which involved working with parole and probation officers to get names and Census Day addresses of parolees and probationers and add them to the census had they not already been counted; (4) the housing coverage check, in which the Census Bureau recanvassed selected blocks based on evidence brought to its attention by the automated management information system; and (5) the postcensus phase of the local government review program.

Recheck of Vacant Housing Units and Those Identified as Uninhabitable or Nonexistent

During the follow-up of nonrespondents by enumerators in May through July, some housing units were identified as vacant or uninhabitable; some addresses were added to the address control file. Each of these units was rechecked by another enumerator in July or August.

Of the approximate 8 million vacancies, the recheck showed 7.6 percent had been occupied as of Census Day, April 1. Their occupants were enumerated at the time of the recheck. This added about 1.6 million persons to the count. Of the approximate 2.9 million units previously identified as uninhabitable or nonexistent, 5.4 percent were reinstated as occupied April 1. These conversions added almost one-half million persons to the count.

"Were You Counted?" Campaign

After the primary data collection, the Census Bureau initiated a procedure to give anyone who thought he/she had been missed the opportunity to fill out publicly available forms or call toll-free 800 numbers that operated in English, Spanish, and six Asian languages. Communities, the media, and many of the 56,000 community-based organizations that had helped initially promote answering the census were encouraged to conduct "Were You

Counted?" campaigns, reproduce census-designed forms or promote call-ins to the 800 numbers. The purpose of th[illegible]paign was to give a second c[illegible]o those who might initially have av[illegible]d being counted, or to reach persons not part of the principal family in a household who might not have been listed on the household questionnaire. Initially, the Census Bureau planned to end the campaign by June 30, 1990, but because so many organizations participated, the toll-free numbers were held open until September 30.

In all, about 400,000 "Were You Counted?" calls or forms came into the Census Bureau. Although the majority of these proved to be persons who had already been counted, the forms did add over 200,000 persons to the census.

Parolee and Probationer Count Check

Research had suggested that a group with a high probability of having been missed in prior censuses were those on parole or probation, a group consisting disproportionately of young males. Thus, in February 1990 the Census Bureau sent letters to the governors and heads of correction departments in each state and the District of Columbia asking them to participate in a program to get parolees and probationers counted. Each was asked to name a liaison to handle the program. Each liaison was sent special individual forms to dis[illegible] to their parole and probation offi[illegible]vho in turn were to distribute them [illegible] those under their jurisdiction.

The response rate for the program was disappointingly low—so low in fact, that the Census Bureau sent enumerators to work with parole and probation officers to complete a form for each parolee/probationer with a verified April 1 address. As a result of this activity, it is estimated over 400,000 persons were added to the census.

Housing Coverage Check

With a computerized census that captured questionnaire data as returns came in, it was possible to make additional accuracy checks not possible in prior censuses. In August of 1990, the Census Bureau searched its data bases to identify any blocks or communities for indications of a low count. While the census was still in progress there was time for a further canvass to make corrections. Population and housing counts, which had accrued thus far for the 39,189 units of local governments, were compared with 1980 counts and recent population estimates. The Census Bureau looked at its data on areas of new construction for possible missed new subdivisions. It also searched to see[illegible]"Were You Counted?" forms showed any pockets of housing that might have been missed. It looked at media reports or local complaints of missed buildings or blocks. Based on these data searches, the Census Bureau decided to recanvass blocks where problems might exist. These blocks represented 15 percent of the Nation's housing units.

Postcensus Local Government Review

39,189 units of local government were sent housing counts and group quarters counts, accrued as of mid-August, to compare with local data. (New updated maps for the communities had already been sent to them in July). Governments were given 15 working days in which to challenge the housing unit or group quarters count for any block. The feedback from local governments was varied. Many took the counts to be final, although the Vacancy Recheck, the Housing Coverage Check—in fact all of the coverage improvement projects done after the primary data collection—were still in progress. All in all, 17 percent of local governments, including all of the 51 largest cities, challenged some blocks, and eight cities challenged over 2,000 blocks. Cities that challenged more than 2,000 blocks in Postcensus Local Review were Atlanta, Boston, Chicago, Detroit, Honolulu, Los Angeles, New York, and Philadelphia.

The recanvass generated by the Housing Coverage Check and Local Government Review yielded new housing units that added over 300,000 persons to the final census count.

The 1990 Post-Enumeration Survey (PES)

Background

The Census Bureau used two major programs to measure coverage for the 1990 census. The first was the Post-Enumeration Survey (PES), which was an independent survey taken after the census and then compared to the census to attempt to measure coverage error in the census. The second program was Demographic Analysis (DA). DA produced an independent estimate of total population by combining information from various sources of administrative data. The process included using historical data on births, deaths, and legal immigration combined with estimates of emigration, undocumented immigration, and Medicare information. Estimates of total population from DA were then compared with census counts to get an estimate of coverage error.

Summary

The PES was a check of the census but not a recount. After the census, interviewers returned to the field to identify all persons living in the sample of blocks at the time of the PES. During the interview, the interviewer asked where each person was living on Census Day—April 1, 1990. This information was then matched to actual census questionnaires. Most people on the PES questionnaires matched to the census. Some did not, and these are the people estimated to have been missed in the actual census. This part of the PES was called the P-sample. People estimated to be missed based on the P-sample were estimated gross omissions in the census.

People can also be included in the census erroneously. An erroneous census enumeration, for example, could be a child born after April 1, 1990, a person who died before April 1, or a college student away from home who was enumerated at his or her parents' address instead of being correctly enumerated at his or her college. Erroneous enumerations also include persons counted twice in the census. Gross erroneous inclusions in the census were measured in the same blocks as the PES and were called the E-sample.

The data on gross erroneous inclusions and gross erroneous omissions were used to produce an estimate of the net undercount or net overcount of the population in the census. This process is described in the following paragraphs. [1]

Selecting the Sample (Sample Design)

The census attempted to cover all people and was conducted in all blocks. The PES was a sample. The PES sample was selected in stages. First a random sample of blocks was chosen. Within sample blocks, all housing units were interviewed. Within an interviewed housing unit, a PES interview was conducted for each person.

Since the PES was a sample, if total population estimates were to be calculated based on it, the results had to be generalized to other people not living in sample blocks. One statistical method to improve the accuracy of this generalization process was to classify sample cases into groups (called post-strata) such that within a group, people were as alike as possible with regard to their propensity to be undercounted. Ancillary evidence indicates that undercoverage is worse for males than

[1] For a more detailed discussion of PES see Howard Hogan, "The 1990 Post-Enumeration Survey: An Overview," a paper presented at the American Statistical Association in August 1990.

females; for minorities than non-minorities; for renters than owners, etc. Therefore, these types of characteristics were used to define the post-strata. The Bureau did not know which post-stratum to assign a person to until after the PES interview was conducted. To help insure an appropriate sample size by post-stratum, the blocks in the U.S. were stratified by similar characteristics before selecting the sample blocks from them.

All blocks in the United States were assigned to one of 101 strata. The strata were defined by geography, city size, racial composition, and percent renter. A representative set of blocks was selected from each stratum. A separate sampling stratum was defined for American Indian Reservations.

Persons living in institutions were excluded from the PES, as were military personnel living in barracks, people living in remote rural Alaska, persons in emergency shelters and persons who had no formal shelter. For each of these categories, it was unreasonable to expect to be able to conduct an independent interview in July and match them to their April 1 location.

The eventual PES sample consisted of about 168,794 housing units in 5,290 block clusters that included 12,124 blocks. (See attachment 1, "PES Sample Size by State.")

The sample was designed to achieve a .7 percent coefficient of variation. That is, the level of sampling error was expected to be .7 percent of the level of estimated undercount or overcount. So for example, if the PES estimated the undercount to be 5 percent, it was expected that the sampling error (or margin of error) on that estimate would be .35 percent. In practice, the sampling error was, on average, 1.7 times more than anticipated by the sample design.

Listing and Enumerating

In February 1990, permanent interviewers of the Census Bureau visited each of the sample blocks to list all housing units they contained. To preserve independence, none of the temporary enumerators hired to take the 1990 census was used for this operation; nor was the listing conducted out of the temporary census offices. To maintain independence, the Census Bureau did not want anyone to know where a PES sample block was so that it would be treated differently during the census.

After the completion of the 1990 census follow-up of those housing units that did not return a questionnaire (called nonresponse follow-up), a set of PES enumerators interviewed persons at households in the PES sample blocks. Although this interviewing drew from enumerators who had worked on 1990 census follow-up, steps were taken to preserve independence, such as not allowing an enumerator to work in a block in the PES that he or she had worked in during the census.

The interviewers determined who was living in each housing unit, obtained their characteristics, and asked where they lived on April 1, 1990, Census Day. The PES interviewing began nearly 3 months after Census Day. Many people had moved during that time. In order to determine whether they were enumerated in the census, the Bureau needed to know where they lived on Census Day and, thus, enumerators asked a series of probing questions to determine occupants' Census Day addresses.

There was a quality assurance program for the interviewing phase to ensure that the interviewers really visited the household and that the people listed were indeed real. If interviewers made up people, they would not match to the census and would inflate the undercount rate.

Matching

The next step was to match the persons enumerated during the PES (the P-sample) to the census. The matching operation was the first step in determining whether persons in the P-sample were enumerated by the census or missed. Basically those persons in the P-sample matched to the census were considered to have been enumerated; those nonmatched were considered to have been missed.

Matching was carried out in four stages. It involved an initial stage of computer matching followed by two stages of clerical matching to attempt to resolve cases that the computer could not match. The two stages of clerical matching were differentiated by the level of skill and judgment required to establish a match.

Those persons in the P-sample not matched to the census by computer and the first two stages of clerical matching were assigned for a follow-up interview, if it was determined that additional information was necessary to establish whether a match to the census was appropriate. An additional fourth stage of clerical matching was then conducted that allowed the more skilled clerical matchers to use the information from the follow-up interview to establish additional matches.

First, the matching classified people as included in the census only if they were counted at the address where they should have been counted, according to the information they provided. This concept was called "correct address" matching. For example, census rules required that a college student be enumerated at the university dormitory, not at his/her parents' home. The PES counted the student as "enumerated" only if he/she was counted at the university. If he/she was not counted at the university, then the student was classified as "omitted" even if he/she were counted at home. In order for the estimation to work out, the enumeration at home was classified as erroneous and subtracted from the census. So in this example, there would have been one omission (at the university) and one erroneous enumeration (at home). The two netted out in the aggregate. The decision to use "correct address" matching was not lightly taken. Indeed, some earlier tests used "any address" matching, i.e., attempting to search all reported addresses. Either approach has advantages and disadvantages.

The second concept was that of the search area. If a person reported that he lived at a given address, then the matching classified him as correctly enumerated if he was counted anywhere in the block. It also classified him as correctly enumerated if he was counted in a surrounding block. There was a limit to how far the matching process could search. If a census computer operation coded the address across town, for example NW vs. SE, the matching did not search there and did not find the person. The matching counted him/her as missed. To balance, the system had to count the other enumeration as erroneous, because it was outside the defined search area.

A final concept was the idea of "sufficient information for matching." When a match was found, it was easy to say that the case was enumerated. When no match was found, it did not necessarily prove that the person was not enumerated, but merely that the search had not been conducted in the correct place. A further review of the case might have shown that there was "insufficient information," leading to its being classified as "unresolved." Rules that classify cases as "sufficient information for matching" were applied before the matching begins. These rules were designed so that for matches there was confidence that the person was correctly enumerated and, equally important, for non-matches, there was confidence that the person was omitted. This approach leads to a somewhat higher "unresolved" rate, but presumably to more accurate overall results.

The accuracy and consistency of the matching process were central to the PES process. Too many matches would have decreased the estimate of population, too few would have increased it. Matching errors would [illegible] distorted the estimated population [illegible]ation if they differed by post-st[illegible]a. The rules were developed over a decade of research. The multiple levels of matching were designed to ensure that the rules were applied consistently between clerks and between offices.

The E-sample, those persons in the PES blocks who were enumerated in the census, was examined to determine if they were correctly enumerated. E-sample persons were matched back into the census to determine if they were enumerated more than once (duplicates). E-sample persons who were matched to the P-sample were assumed to be correctly enumerated (except for duplicate census enumerations). The remaining E-sample persons who were not matched to the P-sample were potential candidates for erroneous enumerations. These unmatched census persons were also included in the PES follow-up operation described above. The follow-up interviewers determined the enumeration status of those persons; that is, if they were correctly enumerated and simply not in the P-sample or if they were erroneously enumerated.

Errors in measuring census erroneous [illegible]erations have almost as much [illegible]n the final estimate of net [illegible]count as errors in measuring census omissions. Reinterview and rematch studies were used to measure the error that the PES makes in measuring census erroneous enumerations and the effects of these errors on the PES estimates.

In processing the E-sample, it was important to include all census enumerations, especially those conducted long after April 1. Common sense and the results from 1980 both indicated that these were more likely to be erroneous than those done on or near April 1. Because of this, there was a special operation to process census enumerations that were enumerated late in the census process. This operation presented special challenges in merging the data with the results of the earlier operation and completing the processing in time.

A final matching and reconciliation operation took place at the conclusion of the PES follow-up. This included the fourth stage of clerical matching for the P-sample and a determination of whether persons in the E-sample were correctly or erroneously enumerated. An important aspect of this operation was that situations arose where correct [illegible] status for persons in the P-sample, or correct enumeration status for persons in the E-sample, could not be determined. This situation occurred because the initial interview was inconclusive or because an incomplete interview was obtained during the follow-up.

Imputation and Dual System Estimation

A final PES file was created that reflected the results of the operations described above. This file included the characteristics of each person in the P-sample and the E-sample. The file also included the match status for persons in the P-sample and the enumeration status (correct or erroneous) for persons in the E-sample. As the final file was prepared, computer editing or imputation was performed to correct, insofar as possible, for missing or contradictory data. A critical aspect of imputation involved the estimation of a final match status for those persons whose match status could not otherwise be resolved. The estimation of match status was very critical. For example, mistakes in the PES matching process, which incorrectly identified persons as not counted in the census (nonmatches), erroneously overstated the estimated undercount and vice versa.

The data in the final PES file were then summarized and incorporated with data from the full census to produce dual system (PES and census) estimates (DSE's) of total population. The DSE's were produced for unique estimation strata (or groupings of persons described below). The dual system estimator is explained more fully in Hogan's document cited above. Essentially it involves estimating how many people were (1) in the PES and in the census, (2) in the PES and out of the census, (3) in the census but not in the PES, and (4) in neither the census nor PES.

The dual system model conceptualized each person as either in or not in the census enumeration, as well as either in or not in the PES. Each person was classified according to the following tableau where the subscripts denote row and column and the stars indicate summing over the entire row/column. N_{**} denotes the entire population.

ENUMERATION

PES	Total	In	Out
Total	N_{**}	N_{*1}	N_{*2}
In	N_{1*}	N_{11}	N_{12}
Out	N_{2*}	N_{21}	N_{22}

All cells were conceptually observable except for N_{22}, and of course any of the marginal totals that include N_{22}. The cell N_{22} (often called the 4[th] cell) was an estimate of people missed in both the census and the PES. Even though not directly observable, the DSE of total population included an estimate of people in the 4[th] cell. The DSE of total population was based on several assumptions. If the PES was an (approximately) unbiased sample of the whole population, then an (approximately) unbiased estimate of N_{**} could be made by noting that the ratio of those in the PES and in the census to the total in the PES should have been the same as the ratio of the total in the census to the total population. Algebraically:

$$N_{11}/N_{1*} = N_{*1}/N_{**}$$

Then solve for the total population:

$$N_{**} = (N_{*1} \cdot N_{1*})/N_{11}$$

This is the dual system estimator of total population.

DSE's were prepared in each of 1,392 post-strata (see next section for a description). Knowing the undercount or overcount rate for each of the groups was important for estimating the net undercount at the local level. It was acceptable for both the PES and the census to have different coverage rates for different post-strata. However, if within a post-stratum, there were sub-groups where both the PES and the census had significantly lower coverage, then the DSE would have been biased.

Another type of bias would have arisen if being enumerated in the census affected the person's response to the PES, or being in the PES affected the person's response to the census enumeration. This would be the case if the PES interviewer and the enumerator compared notes, or if a person refused to cooperate in the census because he had been recently interviewed in PES. The design sought to minimize this effect by conducting the PES after most of the census operations were completed and by conducting the PES out of the Regional Census Centers rather than out of the local District Offices that conducted the enumeration.

Post-Strata

Using the match status and key data, such as age, race, and sex for each person in the sample, the Bureau prepared DSE's of the total population for each of 1,392 groupings of people (post-strata). The reason for forming the post-strata was to group persons who had similar chances of being enumerated in the census. The post-strata were defined by census division, geographic subdivisions such as central cities of large metropolitan statistical areas, whether the person was the owner or renter of the housing unit, race, age, and sex. Each person in the PES sample belonged in one of the unique post-strata. A full description of the 1,392 post-strata is shown in attachment 2.

For purposes of illustration, the following are examples of the 1,392 post-strata. One example is a post-stratum which contains Black males, age 20–29, living in rented housing in central cities in the New York primary metropolitan statistical area. A second example is that which contains non-Black non-Hispanic females, age 45–64, living in owned or rented housing in a non-metropolitan place of 10,000 or more population in the Mountain Division. A third example is that which contains Asian males, age 45–64, living in owned or rented housing in metropolitan statistical areas but not in a central city in the Pacific Division. A fourth example is that which contains non-black Hispanic females, age 30–44, living in owned or rented housing in central cities in the Los Angeles-Long Beach primary metropolitan statistical area or other central cities in metropolitan statistical areas in the Pacific Region. As can be seen from these examples, the 1,392 post-strata are very specific.

The Decision on Combining PES and DA Results Before Computing Adjustment Factors

It was expected that the estimate of total population from the PES would be lower than the estimate of total population from DA. That is because there is a tendency for some people to be missed in both the census and the PES. (often referred to as correlation bias.) No such bias exists with DA estimates. For that reason, there was an open decision point about whether or not to "rake" PES estimates to DA estimates before producing adjustment factors.

After examining the information, the Census Bureau decided against trying to combine the results of DA and PES. There were several reasons for the decision. Some of the main ones include:

• The PES estimate of total population was higher than the DA estimate.

• The PES estimate of females was considerably higher than the DA estimate.

• At the point in time the decision had to be made, the DA estimates were preliminary. There was concern that DA estimates might change considerably over time.

• A concern about the quality of certain components of the DA estimates; for example, the estimate of undocumented immigrants.

• The uncertainty about how combining DA estimates might effect the assumptions underlying the DSE system.

Adjustment Factors

The next step in the post-enumeration survey process was to compare the estimated total population for each post-stratum (the dual system estimate or DSE) to the census count to determine a "raw" adjustment factor. For example, if the DSE for a particular post-stratum was 1,050,000 and the census count was 1,000,000, then the adjustment factor was 1.05, reflecting about a 5-percent estimated net undercount of variability. An adjustment factor may be less than one, thus lowering the census count in a post-stratum if an adjustment is applied. This results when there is evidence of an overcount in the post-stratum.

"Smoothing" the Adjustment Factors

The next steps were "smoothing" the variances of these "raw" adjustment factors, "smoothing" the "raw" adjustment factors themselves to reduce sampling variance associated with them, and the production of final adjustment factors incorporating both smoothing steps. Because the PES was a sample, it was subject to sampling error. Sampling error is an estimate of the error associated with taking some of the population (a sample) rather than all of the population (a census). Disaggregating 377,000 PES persons to 1,392 post-strata produced some post-strata with small sample sizes, and therefore, high estimates of sampling error. The process of smoothing the "raw" adjustment factors to create final adjustment factors was a step to minimize the effect of sampling error.

Both "smoothing" steps were based on a multi-variate regression model. The factor smoothing step used observed characteristics that have been known to be correlated with undercount. A regression prediction model "predicted" the adjustment factor for each of the 1,392 post-strata. The final adjustment factor was then a weighted average of the originally observed adjustment factor (called "raw") and the modeled factor (from the regression prediction model.) For a post-stratum with low estimated sampling variance, there was heavy weight on the observed factor; and vice versa. The final adjustment factors by post-stratum are shown in attachment 3.

Small Area Estimation

The final adjustment factors were now ready to be used to produce adjusted counts for every block in the Nation. The PES can only produce "direct" estimates of the total population for relatively large geographic areas (i.e., the 1,392 post-strata). If there is a decision to adjust, however, the adjustment must be applied to each of the Nation's 4 million populated blocks. The Bureau developed a model that takes the adjustment factors produced for each of the 1,392 post-strata areas and uses them to estimate adjustment counts for each block. Since each of the post-strata crosses many blocks, the Bureau based its model on a critical assumption that coverage error is similar for all blocks that a post-stratum crosses.

Here are two examples of how block counts could be changed during this process. Suppose a census block with 200 people had 50 people who fell into a particular post-stratum. An adjustment factor of 1.05 was computed for that post-stratum, so 50 was multiplied 1.05, which comes to 52.5. Since procedures allowed adding only whole persons to a block, either 2 or 3 persons were added, based on a pre-specified procedure, to the persons in that post-stratum for that block. Other groupings of persons in the block in this example also were multiplied by the adjustment factor for the post-stratum into which they fell. Similarly, suppose there were 80 people in another post-stratum in a particular census block, and the adjustment factor was 0.94, indicating an overcount. 80 was multiplied by 0.94, which came to 75.2, so 4 or 5 person records were eliminated from that block.

The Bureau then produced a data file that included enumerated people plus people added (or subtracted) by adjustment. It did this by adding or subtracting "adjustment" persons with characteristics that were imputed from other persons in the same block. The "adjusted" data files could then be used to produce all required census tabulations.

The 1990 Post Enumeration Survey Evaluation Program

The Post Enumeration Survey (PES) was conducted to evaluate the coverage of the 1990 Decennial Census. Twenty evaluation projects were subsequently conducted to evaluate the PES.[2] This report briefly describes the objectives and implementation of these twenty PES evaluation projects.

[2] In this document, studies P-13 and P-14 are discussed as one study each, although each had two parts. Elsewhere, these parts may be discussed separately, which leads to a total of twenty-two studies.

Ten of the sources of potential error in the PES were addressed by the evaluation studies:

'issing Data.
uality of the Reported Census Day Address.
3. Fabrication in the P-sample.
4. Matching Error.
5. Measurement of Erroneous Enumerations.
6. Balancing the Estimates of Gross Overcount and Gross Undercount.
7. Correlation Bias.
8. Small Area Estimation.
9. Late Census Data.
10. Total Error.

Each of these ten potential sources of error are herein described along with the specific PES Evaluation project used to evaluate or estimate that error.

More detailed project descriptions are found in the Project Plans dated July 31, 1990. For more detailed descriptions of the implementation and results of these projects, see the final reports of July, 1991, whose executive summaries can be found in Appendix 3.

1. Missing Data

Both the P- and E-samples contain missing data on enumeration status. The E-sample has cases where the information required to determine whether the person is correctly or erroneously enumerated in the census is not available. The P-sample has cases w e information needed to de ie whether the person is enumerated in the census is not available.

Missing data occur in more than one way. The interviewer may be unable to obtain an interview during the P-sample interview or during the PES follow-up. A P- or E-sample questionnaire may not have all the demographic and housing information to establish correct enumeration status. Finally, even with all the information requested on the questionnaires, circumstances may be so unclear that the enumeration status cannot be resolved or determined.

Missing data on enumeration status were handled in the production PES in three ways: noninterviews to the P-sample interview were handled by a weight adjustment; missing demographic characteristics in the P- and E-samples (such as age or race) were imputed by means of a hot-deck procedure; and unresolved match status cases were handled by a logistic regression technique.

Missing data can affect the estimates of undercount in a number of ways. For example, if the number of imputed correct enumerations is too high, the undercount estimate will be biased upv if the number of imputed matches in the P-sample is too high, the undercount estimate is biased downward.

Project P1: Analysis of Reasonable Alternatives

The analysis was based on applying alternative missing data treatments, such as methods of handling proxy interviews and mover data, applying bootstrap samples and applying other logistic regression methodologies to study the sensitivity of the dual system estimate to the method of imputation of missing data. A narrow range of alternative estimates indicates robustness in the dual system estimates, indicating little uncertainty in the estimates due to missing data.

The following were the principal alternate imputation treatments:

P-sample Proxy Alternative: P-sample follow-up interviews marked as proxies (i.e. completed with nonhousehold member) were recoded to indicate that no interview was obtained during follow-up.

E-sample Proxy Alternative: E-sample follow-up interviews marked as proxies (i.e. completed with nonhousehold member) were recoded to indicate that no interview was obtained during follow-up.

P-sample Mover Alternative: Unresolved P-sample movers were imputed as if they were nonmovers.

1988 Style Logistic Regression Alternative: The 1990 production imputation model is quite different than the model that was used in the 1988 Dress Rehearsal. The 1988 Style Logistic Regression Model consists of several standard logistic regression models as in 1988.

Bootstrap Samples: Three E-sample and three P-sample bootstrap samples were drawn in order to measure the variation in the production dual system estimates given the PES sample of blocks. Each bootstrap consisted of selecting households with replacement within blocks.

Imputation Treatment Combinations: Dual system estimates were computed for imputation treatment combinations. The following treatment combinations were used:

P-sample Proxy and E-sample Proxy
P-sample Proxy and 1988 Style Model
E-sample Proxy and 1988 Style Model
P-sample Proxy, E-sample Proxy, and 1988 Style Model

Project P2: Distribution of Missing Data Rates

This study was based on analysis of the missing data rates observed for the P- and E- samples. The types of missing data of greatest interest are noninterviews for the initial PES interview, and unresolved cases which remain after the PES follow-up.

The objectives of PES evaluation project P2 are to determine the level and distribution of missing data by demographic and geographic breaks and to compare the distributions with the distribution of census undercount (overcount). Hence, the following estimates are examined for P2.

1. Outcome of Interview (PES, PES Follow-up, and PES Evaluations).
2. Proxy Rates (PES, PES Follow-up, and PES Evaluations).
3. Percentage of Item Imputation (Hot-Deck and Logistic Regression).
4. Correlation Between Item Imputation and Census Undercount.

Project P3: Evaluation of Imputation Methodology for Unresolved Match Status Cases

This study was based on a reinterview of a sample of the P- and E-sample cases that were unresolved after the completion of the PES production follow-up. The reinterview also included a sample of the initial PES incomplete interviews. The reinterview was conducted immediately following the final PES matching operation. The reinterview used a probing questionnaire and better quality interviewers. In addition, the reinterview procedure allowed greater opportunity to contact knowledgeable respondents.

The objectives of PES evaluation project P3 are to: (1) provide quantitative information on the effect of the match/enumeration status imputation procedures; (2) examine quantitative measures of the effect of the noninterview adjustment; and (3) examine the characteristics of the household noninterviews. Hence, the following aspects of the PES are evaluated in P3.

1. Match/Enumeration Status Imputation.
2. Converted PES Noninterview Households.
3. PES Noninterview Household Characteristics.

2. Quality of the Reported Census Day Address

Dual system estimation assumes that P-sample respondents can be linked, or matched, correctly to their census day address. This evaluation measures address reporting and the error in the number of people matching a census enumeration due to address reporting error. Census Day was on April 1, 1990. The PES was conducted in July and August, 1990. Thus, some of the respondents had moved between the time the census was conducted and the PES was in the field. However, in spite of probes on the PES interview questionnaire, respondents may fail to report that they moved. This type of error may cause the matching operation to search the census in an area other than where the respondent was enumerated and to assign a nonmatch status to respondents who might have been enumerated.

Project P4: Quality of the Reported Census Day Address—Evaluation Follow-up

An additional reinterview of a sample of P-Sample cases from the production follow-up was conducted. The sample consisted of nonmatches and unresolved P-sample cases in the PES block clusters selected for the evaluation follow-up. Some matches from whole household matched households were subsampled within each cluster. In addition, matches were selected from partially matched households. A specially designed questionnaire with special probes was used by highly skilled enumerators (Census Bureau Field Representatives). The reinterview allowed greater opportunity to contact designated respondents and probe more deeply for census day accuracy of the PES process for identifying movers and the quality of mover address reporting. Therefore, reviewing these results allowed an assessment of the accuracy of the census day address reported in the production PES.

This evaluation is based on a follow-up and reinterview operation that took place immediately following the final PES matching operation. The follow-up operation consisted of a sample of P-sample matched and nonmatched persons who were excluded from the production follow-up. A review of the results of this follow-up addressed the questions concerning the assumptions underlying the rules that were used in determining which cases should be sent for the production follow-up. This operation was done after PES production matching had been concluded.

3. Fabrication in the P-Sample

Interviewers, for whatever reason, may fabricate persons within enumerated housing units. The PES program had an extensive quality control (QC) program that identified and corrected fabrications. However, even with the best of intentions fabrications potentially remain after this operation. Three studies were implemented to address the effect of any uncorrected fabrications that remained in the data set after the quality control operation. The first study (P5a) identifies the residual fabrication by means of the evaluation follow-up and revisit interviews; subsequent matching of these households will identify fabrications. The second study (P5) utilizes the PES field operation quality control records to estimate "upper bound" residual PES fabrications. The third study (P6) provides model-based estimates of fabrications by comparing, at the block level, interviewer nonmatch rates with "nearby" interviewer nonmatch rates. These comparisons provide an indication of the quality of the interviewers work.

Project P5a: Analysis of P-Sample Fabrication From Evaluation Follow-up Data

The evaluation follow-up described for Project P-4, provided estimates of P-sample fabricated persons. These estimated fabrications can be used as independent estimates (from the quality control) of the level of fabrications in the P-sample. In addition, the quality control operations for the PES interviewing were assessed by comparing the estimated residual error rate from quality control records with the estimated fabrication rate from the follow-up.

Project P5: Analysis of PES P-Sample Fabrications From PES Quality Control Data

The data for project P5 comes from the Quality Control operation of the PES interviewing phase. The purpose of the QC check is to confirm that the PES interviewer visited the correct housing unit and conducted the interview according to the survey procedures. The roster of names, ages and census day addresses are all verified during the interview for the QC sample. A P-sample questionnaire fails the QC check when the household roster is incorrect. When an error is detected, all the recent work of the production interviewer undergoes a QC reinterview. Fabricated households discovered as a result of the QC reinterview are not used and correct interviews are obtained. Overall, approximately 35 percent of the P-sample (i.e., 56,000 households) were reinterviewed in the QC operation of the PES interviewing phase through telephone calls and personal visits.

The central problem or assumption of investigation for project P5 is the estimation of the amount of residual (i.e., undetected) fabrication that exists in the P-sample after the QC operation has been concluded. This analysis provides estimates both in terms of households and persons within these households.

Project P6: Fabrication in the P-Sample: Interviewer Effect

The objective of P6 was to gain knowledge about possible undetected fabrication in the PES. Though it is expected that curbstoners make up only a fraction of the PES work force and the quality control detects and eliminates such curbstoning, the potential impact of undetected fabricated data can be serious. This type of error inflates the undercount estimate. In addition, the inflated nonmatch rates are likely differential, i.e., larger for some post-strata than others.

The purpose of this study was to evaluate the quality control procedure implemented in PES to see how effective it was in detecting fabrication. This was done by developing a model to predict the nonmatch rate from the actual nonmatch rate obtained by interviewers working in areas with households of similar demographic characteristics. The assumption underlying the model was the interviewers working in similar areas would have similar nonmatch rates and the deviations from the model would indicate undetected curbstoning. Standardized scores (Z-scores) were computed for each interviewer rather than comparing the absolute differences between the observed and the expected rates. This was done to take into account the size of an interviewer's assignment. Interviewers with large scores differed greatly from the model predication, and were identified as potential curbstoners or poor quality workers. These enumerators were further studied to determine where they had worked and whether they had been detected by the PES QC operation.

4. Matching Error

Errors can occur in the operation where P-sample persons are matched to the original census enumerations. This matching operation was conducted in seven processing offices (PO's). Even though great efforts were made to standardize this operation across all PO's, errors could be relatively concentrated. Two studies were conducted to examine this type of error. The first study (P7) utilized a team of professionals to dependently rematch a subsample of PES block clusters; this operation is referred to as the Matching Error Study. The rematchers had access to the match codes assigned by the PES production matchers, and worked on assignments in PO's other than their home PO where they worked on PES production. The rematch was designed to estimate the net error rate in the assignment of enumeration status in the ...mple and the E-sample. The second ...y (P5) examined PES production ...ality control records. This analysis provides insight into the nature of PES production matching error by examining where differences occur within this multi-tiered operation.

Project P7: Estimates of Clerical Matching Error From the Evaluation

This evaluation was based on a rematch of a subsample of the PES blocks by highly skilled personnel. This project also allowed additional field work as required, when additional information was determined to be necessary to resolve specific cases. The assumption underlying the evaluation is that better training and personnel can detect systematic errors in the matching.

The subsample of blocks included in this evaluation was based on a stratified sample designed to give a higher probability of selection to blocks with potential matching problems. In addition, the highly skilled personnel used for this evaluation were assigned to work in different processing offices, to the extent possible, to minimize redoing blocks that they previously processed.

Project P8: Matching Error—Estimates of ...cal Matching Error in the P-Sample Quality Assurance Results

This evaluation was carried out by comparing the results of the PES matching quality control operation to determine where potential inconsistencies existed.

At the conclusion of the computer matching, the clerical matching proceeds with an initial stage of clerical matching (CMG) followed by a more extensive stage of matching by another group of more qualified special matching group clerks (SMG1). Another special matching group (SMG2) also conducted matching on the same cases as the CMG and SMG1 stages. Discrepancies between the SMG1 and SMG2 are adjudicated by a higher level PES matching technician.

Comparing the differences between the various stages of matching can identify potential areas where matching error can exist. These findings may be of interest in interpreting the results of project P-7.

5. Measurement of Erroneous Enumerations

Some census enumerations are in fact erroneous. The following enumerations are erroneous:

(1) Duplicated persons.

(2) Fictitious persons.

(3) People who died before Census Day.

(4) People who were born after Census Day.

(5) People enumerated outside the search area where they were living on Census Day.

An estimate of erroneous enumerations is needed for the PES-census dual system estimate of the total population. Three studies investigate errors in classifying the enumeration status (correct or erroneous) of the E-sample persons. The first study (P10) utilized the same team of highly skilled professionals as did project P7 to dependently review the PES E-sample production results in a subsample of PES block clusters. This operation was part of the Matching Error Study. The focus was on the errors that occurred during PES production processing involving duplicates and fictitious persons; however, there was also an examination for the above (3), (4), and (5) type errors. The second study (P9a) utilized data collected from the evaluation follow-up interviews. The evaluation follow-up questionnaire was administered by more competent interviewers than was used by PES production. Also, this questionnaire had more probes than the standard PES production follow-up questionnaire. An alternative estimate of erroneous enumerations resulted from this operation. The third study (P9) is a consistency check; an examination of PES E-sample cross-tabulations provides evidence as to whether a particular type of error in classifying enumeration status is present in the data.

Project P10: Accurate Measurement of Census Erroneous Enumerations—Clerical Error in Assignment of Census Enumeration Status

This evaluation was conducted as part of the rematch work described for Project P7, Evaluation of Clerical Error in the P-sample matching. The study used the same subsample of PES blocks. The E-sample for these blocks underwent the intensive review by highly skilled matchers. This work was supplemented by the reinterview described for Project P9a. The objective was to determine whether the production matching operations are correctly classifying census erroneous enumerations.

The combination of both of these projects—P7 and P10- is referred to as the Matching Error Study (MES).

Project P9a: Accurate Measurement of Census Erroneous Enumerations—Evaluation Follow-up

A sample of E-sample cases was sent for a PES evaluation field follow-up to determine whether a person was correctly enumerated in the Census. The sample included both E-sample cases where an interview was obtained and those where a follow-up interview was not completed. The follow-up reinterview was conducted with more experienced enumerators using a more probing questionnaire. In addition, the follow-up allows greater opportunity to contact a respondent and obtain a complete interview. This same evaluation follow-up was used as part of Project P7 and Project P4. The completed evaluation follow-up interview was clerically matched back to the census to assess the accuracy of the PES production procedure in classifying a persons enumeration status.

Project P9: Accurate Measurement of Census Erroneous Enumeration—Consistency Checks

This evaluation was based on examining a variety of cross tabulations prepared from the PES E-sample for each evaluation stratum. Data such as the following was cross-tabulated:

(1) Enumeration status (correct enumeration, erroneous enumeration).

(2) Type of respondent (original census residents, current residents, neighbors, other proxies).

(3) Source of census enumeration (mailback, enumerator return).

(4) Age group.

(5) Enumeration status of other household members (whole household erroneously enumerated, partial household erroneously enumerated).

The cross tabulations were examined to assess whether the pattern of erroneous enumerations was consistent with previous experience and research findings. Unexplainable discrepancies in the erroneous enumerations were considered as potential indications that the PES process incorrectly measured erroneous enumerations.

6. Balancing the Estimates of Gross Overcount and Gross Undercount

Because of the limited search area that is used to estimate P-sample nonmatches and E-sample erroneous enumerations, balancing error can occur. There was no plan to obtain a direct estimate of this type of error. The components of balancing error are included in the measures of errors that are produced from other studies such as P-7 and P-10 (matching error studies)

Project 11: Balancing Error Evaluation—Percentage of Matches Found Outside Sample Blocks

This evaluation used supplementary information to assess whether balancing is an issue in the performance of PES. Inconsistencies found are indications of potential failure of balancing and should be indications of which of the evaluation studies should reflect these errors. The P-sample match rates for the PES blocks and surrounding blocks were compared with the rates at which E-sample persons are found to be in the PES blocks and in the surrounding blocks. These rates should be the about the same. Differences found were evaluated using the results of the evaluation follow-up.

The rate at which movers matched in the blocks to which they were geocoded was also studied. These rates should be consistent with the corresponding rates for the P-sample nonmovers in the same post-strata.

7. Correlation Bias

The dual system estimation used for the PES is based on several independence assumptions. Two that are of particular interest are homogeneity and causality. The homogeneity assumption requires that everyone has the same probability of inclusion in both the P-sample and the census within the same post-stratum. Failure of the homogeneity assumption usually is seen in an understatement of the undercount for a population group (such as Black males). The causality assumption requires that inclusion in the census does not influence inclusion in the P-sample or vice versa.

Two studies were directed at studying the adequacy of the homogeneity assumption. The first study (P13) compares the dual system estimates with demographic analysis to obtain an estimate of correlation bias at the national level. The second study (P17) is qualitative in nature, and compares the PES dual system estimates, the individual P- and E-samples, and demographic analysis to determine if inconsistencies exist that could indicate the presence of correlation bias due to failure of the homogeneity assumption.

The causality assumption is investigated by two qualitative studies (P14 a and b). The first of these studies pairs non-PES blocks with similar PES blocks and compares characteristics. There should be no difference between these blocks except for the random variation introduced by sampling. The second study uses a debriefing of field interviewers to assess the potential for correlation bias.

Project P13: Use of Alternative Dual System Estimators to Measure Correlation Bias

Alternative dual system estimators were developed using information from demographic analysis to try to address the problem of correlation bias due to failure of the homogeneity assumption—when people missed by the census are more likely to be missed by the PES than those included in the census and vice-versa. This was done by using demographic analysis sex ratios (the ratio of males to females) and the PES dual system estimates for females to create an alternative estimate for males. The DSE for females was multiplied by the sex ratio appropriate for each PES age group. By comparing these alternative estimates for males with the PES dual system estimates for males gives an estimate of correlation bias at the national level. The estimated correlation bias was then allocated to the individual PES male post-strata proportional to P-sample non-matches. This permitted estimates of correlation bias to be produced at the individual post-stratum level.

Project P17 Internal Consistency of Estimates

This study has two objectives: (1) to evaluate the reasonableness of the age sex distribution in the census and PES estimates and (2) to compare the PES and demographic analysis (DA) estimates of undercount to make some assessment of the accuracy of the PES estimates. For these purposes, sex ratios and information on undercount rates from the PES and DA were used. Sex ratio are used to evaluate if overall results on sex distribution are reasonable. Because demographic analysis estimates are available at the national level only, most comparison are limited to analyzing data for the U.S. by race black and non-black.

Project P14 Independence of the Census and P-Sample, Comparison of Blocks

The analysis for this project is directed at assessing the existence of correlation bias due to failure of the causality assumption:

The probability of an individual being included in the P-sample is not altered by inclusion in the census, and the probability of being included in the census is not altered by inclusion in the P-sample.

Several steps were implemented to study the existence of correlation bias. First, a sample of PES blocks paired with comparable non-PES blocks was drawn. The sample was selected by type of enumeration area (TEA) in order to do analyses isolating these groups. Each type of enumeration was analyzed as a separate data set since the timing of the PES and census operations were different across areas. Therefore, any PES effects on the census would be different for each TEA and should be tested using separate data sets.

The difference from PES blocks and non-PES blocks were the focus of the tests. For each block, relevant data were extracted from the final census files in January, 1991 and aggregated from person records to block level records. The preliminary variables were organized a priori into groups: block size, population coverage, housing unit status, mailback, field response, and edit & quality. The data were tested for relevance, completeness, and redundancy.

8. Small Area Estimation

Project P12: Evaluation of the Synthetic Assumption

Synthetic adjustment is used in the PES to "carry down" the estimated adjustment factors to the census counts in each post stratum. This synthetic adjustment assumes that the probability of being missed by the census is constant for each person within the post-stratum.

The coverage error may vary substantially within the PES strata although the post strata were drawn so as to be homogeneous with respect to expected coverage error. The goal of this study is to verify that the assumption underlying the synthetic adjustment is valid.

The analysis was based on studying the homogeneity of several different block level statistics. Three different types of analysis were conducted. First the distributions of census characteristics thought to be highly correlated with coverage error (e.g., mail return rate) were examined. Secondly, the distribution of the components of coverage error at the block level were studied. These components were erroneous enumeration rates and P-sample nonmatch rates. Finally, the production smoothing model was used to predict a block level adjustment factor for the same sample of blocks used for the first analysis.

The analysis concentrated on determining whether the block level statistics clustered unusually by state within the PES post-strata. Further analysis to examine clustering at other levels such as place and county remains to be carried out.

9. Late Late Census Data

Project P18: Evaluation of Late Late Census Data

Census data capture was completed after the completion of the last planned PES matching operation which was Late Census Data matching. A small amount of changes to census data (census additions, deletions and updated person data) resulted from the late census data capture activities. A portion of these changes were included into the PES results through the Late Late Census Data (LLCD) matching operation. The remainder of these late census data changes were not processed due to time constraints, and were not included in the PES results. The Evaluation of Late Late Census Data (Project 18) examines the effect that the late census data changes not included in the PES have on the PES estimates of undercount. The remaining late-late census data were processed to determine the effect that this would have had on the dual system estimates.

10. Total Error

Project 16: Total Error in PES Estimates for Evaluation Post Strata

The dual system estimator used in the estimation for the PES is known to be subject to various components of nonsampling error, in addition to sampling error. The PES evaluation program includes studies that provide direct measures of error due to nonsampling and sampling error components. These errors combine in the dual system estimator model to cause differences from population counts that would be attained under an error-free program. The difference between the PES estimate and the error-free count is referred to as the total error.

Project P16 evaluates both the components of error and the total error in the PES estimates for the 13 evaluation post strata. The components of error are response correlation bias (also called model bias), matching error, quality of reported Census Day address, fabrication in the P-sample, processing error in the E-sample, data collection error in the E-sample, error in balancing the estimates of the gross overcount and the gross undercount missing data (imputation error), sampling variance, and ratio estimator bias.

The evaluation of the total error assesses the overall accuracy of the PES estimates of population size and the census undercount rate. A synthesis of the components errors provides estimates of the bias and variance. This analysis then assesses the combined effect of the errors on the PES estimate of the undercount rate. The estimates of the mean and variance of the distributions of the component errors are based on the conclusions drawn from the various evaluation studies. The simulation method produced an estimate of the bias and variance of the estimated undercount rate.

The results of the total error model were also used in a loss function analysis to assess the accuracy of the distributions of population across states, places, and counties for the adjusted and unadjusted census. This analysis was carried out by forming target populations from the results of the total error work. The biases measured by the PES evaluations were incorporated into PES dual system estimates to produce corrected estimates of the population. These corrected estimates were designated as the target populations. The adjusted and unadjusted census population distributions were compared to the target population distributions using several loss functions. The comparisons were conducted at the state level and at the place and county level for the following size categories:

Places under 25,000 population.
Places of between 25,000 and 50,000 population.
Places of size over 50,000.
Counties under 200,000.
Counties larger than 200,000.

In addition, results were also produced for places and counties over 100,000 population.

Demographic Analysis

The Census Bureau's companion coverage measurement program to the PES was demographic analysis. The demographic coverage estimates could only be used to evaluate the completeness of coverage of the 1990 census at a national level and only for race (Black/Non-Black), sex, and age groups. Demographic analysis could not provide even reasonably reliable coverage estimates for the Hispanic, Asian/Pacific Islander, or American Indian/Native Alaskan populations because these characteristics have not always been recorded on birth and death certificates; nor can the demographic method provide direct estimates of the resident population at the State or substate level. However, the PES measured under or overcounts of these groups. The demographic coverage estimates were compared to the post-enumeration survey coverage estimates to assess the overall consistency of the two sets of estimates at the national level.

Demographic analysis uses historical data on births, deaths, and legal immigration; estimates of emigration and undocumented immigration; and Medicare data to develop an independent estimate of the resident population on census day. The estimate is compared with the census count to yield a measure of net census coverage and net undercount. The particular procedure that is used to estimate coverage nationally in 1990 for the various demographic subgroups depends primarily on the nature and availability of the required demographic data. Birth and death records are available for the entire United States from 1933 on for developing estimates of population at ages under 57 in 1990. In estimating births for each year, the Bureau added to the number of registered births an estimate of underregistration. Underregistration was estimated based on tests conducted in 1940, 1950, and 1964–1968. If the estimates of underregistration are off, they could have a significant effect on undercount estimates because birth data are by far the largest component in estimating the population through demographic analysis. In fact, in producing the demographic estimates of population for 1990 the Bureau revised the estimates for certain Black birth cohorts to account for biases that recent research identified in the birth registration test result of 1940.

National birth and death records are not available before 1933, so the Bureau had to find other ways to estimate the population size of these cohorts in 1990 (ages 55 and over were estimated). For the population 65 and over, administrative data on aggregate Medicare enrollments for 1990 (adjusted for underenrollment) are used to estimate population and net coverage. For the Non-black population aged 55 to 64 in 1990, the estimates of population are based primarily on national birth estimates for 1925–1934 developed by Whelpton. For the Black population aged 55 to 64 in 1990, the estimates of population are based on revisions of estimates for the cohort in 1960 developed by Coale and Rives.

In addition to subtracting deaths, the estimates of births described above are augmented to account for change due to immigration, emigration, and net international movement abroad of citizens (including the Armed Forces and Puerto Rican migrants). The various components of net migration vary significantly in their completeness and quality. The United States does not keep emigration records. Therefore, an estimate had to be made of those who have left the country. While the United States does have good records of legal immigration, there is no accurate estimate of illegal immigration—the most elusive demographic component of population change. The Census Bureau has developed a preliminary estimate for undocumented residents in 1990 based on analysis of survey data and administrative records of the Immigration and Naturalization Service (INS). The INS now collects different information than it did prior to 1980. Recent immigration reform further complicated the effort to estimate legal immigration and undocumented residents. Although the legislative reform allowed many undocumented aliens to receive amnesty, some of these persons may not actually reside in the United States.

It should be noted that before the demographic estimates of population for race groups are compared to the census to calculate the net undercount, the race categories of the census counts must be "modified" so that they are consistent with the race categories of the historical demographic estimates. Specifically, 9.8 million persons in the 1990 census (mostly of Hispanic origin) reported their race in the "Other race-not specified" category, a category not included in the demographic estimates. This modification added 497,000 persons to the census count for Blacks. Also, the age categories of the 1990 census counts have been "modified" so they are consistent with the April 1, 1990 time reference of the demographic estimates.

It is important to emphasize that results of demographic analysis are not exact but are estimates. To a large extent, they were based on assumptions and best professional judgment. As in the PES, the Bureau tried to estimate potential error in the data produced by demographic analysis. To estimate that overall error, the Bureau conducted 11 detailed demographic analysis evaluation studies to find out as much as possible about each possible source of error—the specific projects are identified in Table 1. Based on these studies, the Bureau developed a range of error around the demographic analysis estimates. Since these evaluation projects and the demographic error model represent an evaluation program new for the 1990 census, the assessments of potential error are subject to change and improvement over time just as the basic demographic estimates of coverage have been.

Table 1.—The Eleven Demographic Analysis Evaluation Projects

D1 ...... Error in Birth Underregistration Completeness Estimates.
D2 ...... Uncertainty in Estimates of Undocumented Aliens.
D3 ...... Uncertainty in Estimated White Births, 1915–1935.
D4 ...... Uncertainty in Estimated Black Births, 1915–1935.
D5 ...... Robustness of Estimated Number of Emigrants.
D6 ...... Robustness of Estimates of the Population 65 and Older.
D7 ...... Uncertainty Measures for Other Components.
D8 ...... Uncertainty of Models to Translate 1990 Census Concepts into Historical Racial Classifications.
D9 ...... Inconsistencies in Race Classifications of the Demographic Estimates and the Census.
D10 ...... Differences Between Preliminary and Final Demographic Estimates.
D11 ...... Total Error in the Demographic Estimates.

Attachment 1

PES SAMPLE SIZE BY STATE (P-SAMPLE)

State names	Blocks	Clusters	Housing units
Alabama	280	168	4,706
Alaska	27	16	846
Arizona	569	115	5,046
Arkansas	161	77	2,230
California	652	390	13,013
Colorado	401	101	3,290
Connecticut	74	55	1,816
Delaware	19	12	460
District of Columbia	22	18	657
Florida	296	198	5,973
Georgia	189	112	3,320
Hawaii	49	19	599
Idaho	226	51	1,897
Illinois	300	221	7,553
Indiana	149	92	2,540
Iowa	179	66	2,491
Kansas	264	74	2,188
Kentucky	177	107	3,116
Louisiana	165	105	3,481
Maine	216	67	2,292
Maryland	72	56	2,162
Massachusetts	162	107	3,185
Michigan	232	152	4,959
Minnesota	256	99	3,196
Mississippi	179	103	2,896
Missouri	215	116	3,369
Montana	409	46	1,755
Nebraska	140	44	1,257
Nevada	96	27	1,195
New Hampshire	118	48	1,967
New Jersey	117	91	2,752
New Mexico	553	68	2,533
New York	520	371	12,210
North Carolina	209	126	3,764
North Dakota	205	19	679
Ohio	216	146	4,491
Oklahoma	271	93	2,737
Oregon	310	83	2,575
Pennsylvania	499	303	8,517
Rhode Island	32	24	832
South Carolina	107	58	1,800
South Dakota	230	18	686

PES SAMPLE SIZE BY STATE (P-SAMPLE)—Continued

State names	Blocks	Clusters	Housing units
Tennessee	243	173	4,858
Texas	845	436	12,807
Utah	212	40	1,351
Vermont	115	28	1,423
Virginia	144	87	2,809
Washington	352	111	3,939
West Virginia	49	31	911
Wisconsin	141	76	2,264
Wyoming	486	28	801
National Total	12,124	6,290	166,794

Attachment 2

1990 Post-Enumeration Survey Post Strata

The 1990 Post-Enumeration Survey (PES) will provide direct estimates for 1392 post strata. The post strata are designed to divide the PES sample blocks into groups which have similar characteristics. This helps the Census Bureau to estimate the coverage of the 1990 decennial census more accurately.

The post strata are defined by census division, area (city, non-city, rural, etc.), race, Hispanic origin, tenure group, sex, and age. Tenure refers to whether housing units are owned or rented. Each post strata is given an eight digit code. The attached document shows 116 post strata and the corresponding first six digits of the post stratum code for each. The last two digits are not delineated on the attachment. They define sex and age group. There are six age group classifications. What follows is an explanation of the post strata coding system:

The first digit of each given eight digit code defines the census division. The nine census divisions and the states in each census division are:

1—New England—Connecticut, Massachusetts, Maine, New Hampshire, Rhode Island, and Vermont
2—Middle Atlantic—New Jersey, New York, and Pennsylvania
3—South Atlantic—Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, and West Virginia
4—East South Central—Alabama, Kentucky, Mississippi, Tennessee
5—West South Central—Arkansas, Louisiana, Oklahoma, and Texas
6—East North Central—Illinois, Indiana, Michigan, Ohio, and Wisconsin
7—West North Central—Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota
8—Mountain—Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, and Wyoming

9— Pacific—Alaska, California, Hawaii, Oregon, and Washington

Within each census division, the geographic areas are divided by type of area. There are nine possible type of area codes:

0—Central cities in explicitly named PMSAs (see description below)
1—Central cities in large metropolitan areas (Type I MAs)
2—Central cities in small metropolitan areas (Type II MAs)
3—Central cities in a metropolitan area regardless of size
4—Non-central city areas in the New York PMSA
5—Non-central city areas in large metropolitan areas (Type I MAs)
6—Non-central city areas in small metropolitan areas (Type II MAs)
7—Non-central city areas in metropolitan areas
8—Non-metropolitan areas incorporated places with 10,000 + population
9—Balance of non-metropolitan areas

A PMSA is a Primary Metropolitan Statistical Area. There are four explicitly named PMSAs in the 1990 PES post strata. These PMSAs and the census division in which they are located are:

• The New York City PMSA in the Middle Atlantic division.
• The Houston PMSA plus the Dallas PMSA, plus the Fort Worth PMSA in the West South Central division.
• The Chicago PMSA plus the Detroit PMSA in the East North Central division.
• The Los Angeles-Long Beach PMSA in the Pacific division.

A large metropolitan area (type I MA) is an area whose largest central city has a population of at least 250,000 using the 1990 census person count.

A small metropolitan area (type II MA) is an area which does not have any central cities with a population of 250,000 or more.

The balance of non-metropolitan areas consist of areas which are not included in area type number 8. This would consist primarily of rural areas.

Any post strata can include up to three area types. The area types included in a stratum are delineated in the second to fourth digits of the post strata code. For instance, post strata code 578910 includes area types 7, 8, and 9. But most post strata contain only one area type. If a post stratum has only one area type, the second digit of the post stratum code indicates the area type, and the third and fourth digits are zero. In general, each of the second through fourth digits is filled with a zero from the right if a given geographic area of post stratum contains less than three area types.

The race/hispanic origin is determined by the fifth digit of the post stratum code. The tenure group is determined by the sixth digit of the post stratum code. These three attributes are combined in the coding system. The possible race/hispanic origin groups are: Black, Non-Black Hispanic, Asian-Pacific Islander, American Indian, and Other. A post stratum can consist of more than one race/hispanic origin group. This is reflected in the definitions below. The tenure designation defines whether the persons in the geographic area are owners or renters. Some geographic areas were not divided by tenure. The possible codes for the fifth and sixth digits are:

10—Black (Renter & Owner)
11—Black Renter
12—Black Owner
20—Non-Black Hispanic (Renter & Owner)
21—Non-Black Hispanic Renter
22—Non-Black Hispanic Owner
30—All Other (Renter & Owner)
31—All Other Renter
32—All Other Owner
40—Asian-Pacific Islander (Renter & Owner)
41—Asian-Pacific Islander Renter
42—Asian-Pacific Islander Owner
50—Black and Non-Black Hispanic (Renter & Owner) & Non-Black Non-Asian-Pacific Islander Hispanic
60—American Indian

The seventh digit of the post stratum code defines the sex.

1—Male
2—Female

Within sex there are six age groups, the eighth digit. The age groups are:

1—0–9
2—10–19
3—20–29
4—30–44
5—45–64
6—65+

ATTACHMENT 3.—ADJUSTMENT FACTORS BY POST STRATUM [1]

Stratum code	Factor
09006011	1.166
09006012	1.182
09006013	1.158
09006014	1.197
09006015	1.117
09006016	1.143
09006021	1.130
09006022	1.189
09006023	1.166
09006024	1.071
09006025	1.066
09006026	1.097
13003011	1.001
13003012	0.967
13003013	1.034
13003014	0.984
13003015	0.991
13003016	0.964
13003021	0.969
13003022	0.979
13003023	1.007
13003024	0.976
13003025	0.961
13003026	0.957
13705011	1.068
13705012	1.027
13705013	1.079
13705014	1.066
13705015	1.040
13705016	1.012
13705021	1.047
13705022	1.008
13705023	1.050
13705024	1.041
13705025	1.012
13705026	1.015
17003011	1.020
17003012	0.969
17003013	1.030
17003014	0.990
17003015	1.014
17003016	0.967
17003021	1.016
17003022	0.974
17003023	1.021
17003024	1.007
17003025	0.994
17003026	0.975
18003011	1.025
18003012	0.960
18003013	1.030
18003014	1.026
18003015	1.011
18003016	0.964
18003021	1.007
18003022	0.974
18003023	1.003
18003024	1.006
18003025	1.002
18003026	0.995
19003011	1.022
19003012	1.008
19003013	1.073
19003014	1.028
19003015	1.024
19003016	1.013
19003021	1.017
19003022	1.003
19003023	1.018
19003024	1.013
19003025	0.996
19003026	1.006
20001111	1.111
20001112	1.076
20001113	1.122
20001114	1.102
20001115	1.043
20001116	1.077
20001121	1.112
20001122	1.031
20001123	1.090
20001124	1.114
20001125	1.036
20001126	1.050
20001211	1.022
20001212	0.904
20001213	1.010
20001214	0.990
20001215	0.991
20001216	0.960
20001221	1.065
20001222	0.897
20001223	1.019
20001224	0.969
20001225	0.962
20001226	0.961
20002011	1.050
20002012	0.990
20002013	1.053
20002014	1.018
20002015	1.024

ATTACHMENT 3.—ADJUSTMENT FACTORS BY POST STRATUM [1]—Continued

Stratum code	Factor
20002016	1.002
20002021	0.995
20002022	1.002
20002023	1.033
20002024	1.015
20002025	1.006
20002026	0.994
20003111	0.993
20003112	0.997
20003113	1.113
20003114	1.041
20003115	1.016
20003116	0.964
20003121	1.001
20003122	0.954
20003123	1.054
20003124	1.011
20003125	0.967
20003126	0.935
20003211	0.966
20003212	0.993
20003213	1.013
20003214	1.001
20003215	1.017
20003216	0.954
20003221	1.030
20003222	0.960
20003223	1.017
20003224	1.012
20003225	0.972
20003226	1.002
20004011	1.130
20004012	1.124
20004013	1.156
20004014	1.107
20004015	1.104
20004016	1.095
20004021	1.128
20004022	1.069
20004023	1.130
20004024	1.133
20004025	1.101
20004026	1.061
21001111	1.002
21001112	1.037
21001113	1.126
21001114	1.107
21001115	1.063
21001116	1.033
21001121	1.090
21001122	1.076
21001123	1.127
21001124	1.083
21001125	1.065
21001126	1.035
21001211	1.022
21001212	1.040
21001213	1.029
21001214	0.969
21001215	0.992
21001216	0.968
21001221	1.037
21001222	0.984
21001223	1.010
21001224	0.969
21001225	0.966
21001226	0.969
21003111	1.002
21003112	0.970
21003113	1.034
21003114	1.003
21003115	0.969
21003116	0.984
21003121	0.991
21003122	0.997
21003123	1.001
21003124	1.008
21003125	0.985
21003126	0.959
21003211	1.024
21003212	0.958
21003213	1.013
21003214	1.020
21003215	0.996
21003216	0.982
21003221	1.005
21003222	0.995
21003223	1.048
21003224	0.987
21003225	0.991
21003226	0.979
22001011	1.127
22001012	1.031
22001013	1.129
22001014	1.142
22001015	1.105
22001016	0.957
22001021	1.157
22001022	1.080
22001023	1.140
22001024	1.071
22001025	1.074
22001026	1.058
22003011	0.991
22003012	0.969
22003013	1.037
22003014	1.022
22003015	1.017
22003016	0.975
22003021	1.006
22003022	0.968
22003023	1.002
22003024	0.993
22003025	1.009
22003026	0.974
23002011	1.010
23002012	1.021
23002013	1.071
23002014	1.022
23002015	1.008
23002016	0.972
23002021	1.024
23002022	0.976
23002023	1.008
23002024	1.055
23002025	1.010
23002026	0.995
24003011	1.053
24003012	0.991
24003013	1.020
24003014	1.012
24003015	0.996
24003016	0.961
24003021	1.017
24003022	1.006
24003023	1.057
24003024	0.991
24003025	0.979
24003026	0.978
24505011	1.071
24505012	1.057
24505013	1.115
24505014	1.085
24505015	1.080
24505016	1.080
24505021	1.108
24505022	1.032
24505023	1.083
24505024	1.085
24505025	1.067
24505026	1.014
25003011	1.009
25003012	0.963
25003013	1.037
25003014	1.031
25003015	0.981
25003016	0.971
25003021	1.029
25003022	1.018
25003023	1.033
25003024	1.002
25003025	0.983
25003026	0.973
26003011	1.018
26003012	0.990
26003013	1.040
26003014	0.994
26003015	0.991
26003016	0.964
26003021	1.003
26003022	0.978
26003023	0.999
26003024	1.011
26003025	0.994
26003026	0.984
28003011	1.015
28003012	0.967
28003013	1.017
28003014	1.030
28003015	1.006
28003016	0.991
28003021	1.061
28003022	0.975
28003023	1.016
28003024	0.998
28003025	0.992
28003026	0.984
29003011	1.014
29003012	0.991
29003013	1.042
29003014	1.019
29003015	0.993
29003016	1.001
29003021	1.009
29003022	0.999
29003023	1.041
29003024	1.017
29003025	0.982
29003026	0.986
29005011	1.071
29005012	1.048
29005013	1.067
29005014	1.074
29005015	1.037
29005016	1.033
29005021	1.055
29005022	1.045
29005023	1.054
29005024	1.068
29005025	1.054
29005026	1.039
31001111	1.133
31001112	1.102
31001113	1.106
31001114	1.131
31001115	1.076
31001116	1.066
31001121	1.155
31001122	1.096
31001123	1.105
31001124	1.069
31001125	1.067
31001126	1.037
31001211	1.086
31001212	1.017
31001213	1.030
31001214	1.024
31001215	0.990
31001216	0.991
31001221	1.037
31001222	1.006
31001223	1.027
31001224	0.982
31001225	0.984

ATTACHMENT 3.—ADJUSTMENT FACTORS BY POST STRATUM [1]—Continued

Stratum code	Factor
31001226	0.977
31003111	1.086
31003112	1.038
31003113	1.073
31003114	1.065
31003115	1.047
31003116	0.993
31003121	1.054
31003122	1.055
31003123	1.099
31003124	1.013
31003125	1.011
31003126	0.983
31003211	1.039
31003212	1.035
31003213	1.048
31003214	1.035
31003215	0.983
31003216	0.985
31003221	1.035
31003222	1.031
31003223	1.073
31003224	1.021
31003225	0.979
31003226	1.008
32001011	1.052
32001012	1.035
32001013	1.072
32001014	1.037
32001015	1.015
32001016	1.008
32001021	1.064
32001022	1.028
32001023	1.083
32001024	1.047
32001025	1.003
[illegible]1026	0.981
[illegible]011	1.065
[illegible]012	1.068
[illegible]013	1.080
32003014	1.046
32003015	1.027
32003016	0.986
32003021	1.048
32003022	1.032
32003023	1.039
32003024	1.007
32003025	0.987
32003026	0.998
33002011	1.106
33002012	1.064
33002013	1.101
33002014	1.068
33002015	1.005
33002016	0.985
33002021	1.101
33002022	1.056
33002023	1.091
33002024	1.065
33002025	0.984
33002026	0.964
35001011	1.042
35001012	1.012
35001013	1.034
35001014	1.007
35001015	0.996
35001016	0.990
35001021	1.040
35001022	1.012
35001023	1.045
35001024	1.017
35001025	1.007
35001026	0.986
35003011	1.030
35003012	0.997
35003013	1.032
[illegible]14	1.008
[illegible]15	0.982
35003016	0.985
35003021	1.036
35003022	1.015
35003023	1.035
35003024	0.995
35003025	0.975
35003026	0.983
36001011	1.074
36001012	1.033
36001013	1.034
36001014	1.044
36001015	1.018
36001016	1.003
36001021	1.036
36001022	1.043
36001023	1.051
36001024	1.042
36001025	1.010
36001026	1.001
36003011	1.052
36003012	1.007
36003013	1.039
36003014	1.042
36003015	0.991
36003016	0.992
36003021	1.089
36003022	1.062
36003023	1.038
36003024	1.043
36003025	1.028
36003026	0.994
37892011	1.030
37892012	1.083
37892013	1.133
37892014	1.074
37892015	1.007
37892016	1.017
37892021	1.090
37892022	1.021
37892023	1.068
37892024	1.059
37892025	0.994
37892026	0.971
38001011	1.025
38001012	1.001
38001013	1.023
38001014	1.033
38001015	1.023
38001016	0.984
38001021	1.057
38001022	1.015
38001023	1.048
38001024	1.021
38001025	0.953
38001026	0.963
38003011	1.058
38003012	1.015
38003013	1.088
38003014	1.020
38003015	1.000
38003016	0.981
38003021	1.046
38003022	1.010
38003023	1.026
38003024	1.007
38003025	0.995
38003026	0.979
39001011	1.057
39001012	1.039
39001013	1.021
39001014	1.039
39001015	1.023
39001016	0.981
39001021	1.071
39001022	1.045
39001023	1.045
39001024	0.999
39001025	0.994
39001026	0.979
39003011	1.047
39003012	1.014
39003013	1.063
39003014	1.035
39003015	1.002
39003016	0.994
39003021	1.058
39003022	1.045
39003023	1.060
39003024	1.022
39003025	1.022
39003026	0.997
41003111	1.064
41003112	1.074
41003113	1.056
41003114	1.078
41003115	1.015
41003116	0.989
41003121	1.075
41003122	1.050
41003123	1.042
41003124	1.062
41003125	1.025
41003126	0.982
41003211	1.045
41003212	1.032
41003213	1.042
41003214	1.043
41003215	1.011
41003216	0.994
41003221	1.065
41003222	1.028
41003223	1.084
41003224	1.036
41003225	1.006
41003226	1.001
42003011	1.075
42003012	1.018
42003013	1.072
42003014	1.042
42003015	1.002
42003016	0.996
42003021	1.036
42003022	1.055
42003023	1.058
42003024	1.020
42003025	0.987
42003026	0.975
43005011	1.093
43005012	1.075
43005013	1.090
43005014	1.065
43005015	1.055
43005016	1.009
43005021	1.116
43005022	1.041
43005023	1.083
43005024	1.043
43005025	1.003
43005026	0.987
47003011	1.041
47003012	1.023
47003013	1.043
47003014	1.042
47003015	1.002
47003016	0.987
47003021	1.051
47003022	1.024
47003023	1.050
47003024	1.015
47003025	1.007
47003026	0.990
47895011	1.062
47895012	1.006
47895013	1.020
47895014	1.042
47895015	1.004

ATTACHMENT 3.—ADJUSTMENT FACTORS BY POST STRATUM [1]—Continued

Stratum code	Factor
47895016	0.999
47895021	1.050
47895022	1.011
47895023	1.036
47895024	1.027
47895025	1.004
47895026	0.975
48003011	1.039
48003012	1.029
48003013	1.053
48003014	1.020
48003015	1.003
48003016	0.996
48003021	1.048
48003022	1.017
48003023	1.032
48003024	1.014
48003025	0.988
48003026	1.004
49003011	1.032
49003012	1.021
49003013	1.066
49003014	1.012
49003015	0.990
49003016	0.998
49003021	1.032
49003022	1.016
49003023	1.060
49003024	1.010
49003025	0.967
49003026	1.010
50001011	1.098
50001012	1.061
50001013	1.096
50001014	1.072
50001015	1.042
50001016	1.020
50001021	1.104
50001022	1.057
50001023	1.117
50001024	1.058
50001025	1.022
50001026	0.995
50002011	1.068
50002012	1.044
50002013	1.143
50002014	1.063
50002015	1.014
50002016	0.963
50002021	1.128
50002022	1.079
50002023	1.105
50002024	1.043
50002025	0.992
50002026	0.956
50003111	1.058
50003112	1.050
50003113	1.073
50003114	1.060
50003115	1.035
50003116	1.008
50003121	1.060
50003122	1.043
50003123	1.053
50003124	1.019
50003125	1.028
50003126	0.990
50003211	1.033
50003212	1.004
50003213	1.054
50003214	1.020
50003215	1.017
50003216	0.977
50003221	1.033
50003222	1.019
50003223	1.027
50003224	1.025
50003225	0.997
50003226	0.999
51003111	1.041
51003112	1.036
51003113	1.027
51003114	1.032
51003115	1.014
51003116	0.982
51003121	1.059
51003122	1.039
51003123	1.069
51003124	1.026
51003125	0.999
51003126	0.960
51003211	1.032
51003212	1.014
51003213	1.039
51003214	1.012
51003215	0.994
51003216	0.984
51003221	1.027
51003222	1.011
51003223	1.041
51003224	1.005
51003225	0.994
51003226	0.965
52003011	1.053
52003012	1.034
52003013	1.056
52003014	1.038
52003015	1.003
52003016	0.995
52003021	1.045
52003022	1.017
52003023	1.053
52003024	1.027
52003025	0.994
52003026	0.966
53001011	1.079
53001012	1.034
53001013	1.099
53001014	1.057
53001015	1.032
53001016	1.001
53001021	1.069
53001022	1.047
53001023	1.074
53001024	1.045
53001025	0.969
53001026	0.997
53002011	1.065
53002012	1.037
53002013	1.051
53002014	1.033
53002015	0.974
53002016	0.970
53002021	1.095
53002022	1.023
53002023	1.094
53002024	1.048
53002025	0.971
53002026	0.979
57003011	1.044
57003012	1.043
57003013	1.060
57003014	1.024
57003015	0.999
57003016	0.989
57003021	1.033
57003022	1.032
57003023	1.048
57003024	1.030
57003025	0.993
57003026	0.971
57891011	1.069
57891012	1.018
57891013	1.041
57891014	1.033
57891015	1.004
57891016	0.988
57891021	1.032
57891022	1.003
57891023	1.060
57891024	1.013
57891025	0.992
57891026	0.985
57892011	1.056
57892012	1.068
57892013	1.077
57892014	1.065
57892015	1.047
57892016	1.002
57892021	1.061
57892022	1.047
57892023	1.059
57892024	1.043
57892025	1.011
57892026	1.015
58003011	1.027
58003012	0.994
58003013	1.048
58003014	1.030
58003015	0.999
58003016	0.978
58003021	1.027
58003022	1.021
58003023	1.062
58003024	1.013
58003025	0.993
58003026	0.969
59003011	1.052
59003012	1.032
59003013	1.029
59003014	1.019
59003015	1.014
59003016	0.990
59003021	1.043
59003022	1.016
59003023	1.043
59003024	1.034
59003025	1.006
59003026	0.983
60001111	1.131
60001112	1.031
60001113	1.067
60001114	1.081
60001115	1.042
60001116	0.984
60001121	1.112
60001122	1.061
60001123	1.106
60001124	1.020
60001125	1.004
60001126	0.954
60001211	1.047
60001212	1.001
60001213	1.042
60001214	1.039
60001215	1.015
60001216	1.029
60001221	1.050
60001222	1.040
60001223	1.033
60001224	1.007
60001225	0.994
60001226	0.975
60003111	1.016
60003112	1.021
60003113	1.067
60003114	1.093
60003115	1.106
60003116	1.026
60003121	1.111
60003122	1.052
60003123	1.123
60003124	0.969
60003125	1.017

ATTACHMENT 3.—ADJUSTMENT FACTORS BY POST STRATUM [1]—Continued

Stratum code	Factor
60003126	0.924
60003211	1.021
60003212	1.010
60003213	1.001
60003214	1.027
60003215	1.016
60003216	1.005
60003221	1.047
60003222	1.005
60003223	1.031
60003224	1.000
60003225	1.000
60003226	1.000
60102011	0.958
60102012	1.005
60102013	1.000
60102014	1.026
60102015	0.991
60102016	0.998
60102021	0.993
60102022	0.989
60102023	0.986
60102024	0.957
60102025	0.936
60102026	0.963
61001111	1.042
61001112	1.003
61001113	1.100
61001114	1.052
61001115	1.034
61001116	0.999
61001121	1.066
61001122	1.072
61001123	1.058
61001124	1.047
61001125	0.994
61001126	0.955
61001211	1.091
61001212	0.983
61001213	1.025
61001214	1.026
61001215	1.006
61001216	1.007
61001221	1.045
61001222	1.014
61001223	1.012
61001224	1.000
61001225	0.949
61001226	0.977
61003111	1.119
61003112	0.954
61003113	0.992
61003114	1.070
61003115	1.033
61003116	0.970
61003121	1.030
61003122	0.980
61003123	1.010
61003124	0.999
61003125	0.840
61003126	0.972
61003211	0.957
61003212	0.971
61003213	1.036
61003214	1.021
61003215	0.973
61003216	0.994
61003221	0.986
61003222	0.989
61003223	1.011
61003224	1.003
61003225	1.016
61003226	1.001
62003011	1.033
62003012	0.978
62003013	1.064
62003014	1.019
62003015	1.016

ATTACHMENT 3.—ADJUSTMENT FACTORS BY POST STRATUM [1]—Continued

Stratum code	Factor
62003016	1.020
62003021	1.022
62003022	1.012
62003023	1.032
62003024	1.032
62003025	1.002
62003026	1.008
62705011	1.068
62705012	1.054
62705013	1.080
62705014	1.062
62705015	1.021
62705016	1.006
62705021	1.095
62705022	1.066
62705023	1.074
62705024	1.030
62705025	1.024
62705026	1.002
65003011	1.017
65003012	0.999
65003013	1.030
65003014	1.014
65003015	1.002
65003016	1.003
65003021	1.011
65003022	1.012
65003023	1.008
65003024	0.995
65003025	0.992
65003026	0.997
66003011	1.017
66003012	0.968
66003013	1.023
66003014	1.034
66003015	0.999
66003016	0.972
66003021	1.013
66003022	1.008
66003023	1.000
66003024	1.018
66003025	0.978
66003026	1.002
68003011	1.005
68003012	0.977
68003013	1.028
68003014	1.008
68003015	1.008
68003016	0.997
68003021	1.003
68003022	1.005
68003023	1.019
68003024	0.997
68003025	0.988
68003026	0.987
69003011	0.991
69003012	0.981
69003013	1.019
69003014	0.987
69003015	0.986
69003016	0.997
69003021	0.984
69003022	0.981
69003023	1.014
69003024	0.982
69003025	0.982
69003026	0.995
71003111	1.064
71003112	0.995
71003113	1.107
71003114	1.054
71003115	1.041
71003116	0.997
71003121	1.007
71003122	1.015
71003123	1.012
71003124	0.981
71003125	0.996

ATTACHMENT 3.—ADJUSTMENT FACTORS BY POST STRATUM [1]—Continued

Stratum code	Factor
71003126	0.940
71003211	0.987
71003212	0.980
71003213	1.021
71003214	0.992
71003215	0.996
71003216	0.981
71003221	1.005
71003222	1.026
71003223	1.004
71003224	0.994
71003225	1.000
71003226	0.986
71005011	1.110
71005012	1.030
71005013	1.080
71005014	1.049
71005015	1.037
71005016	1.011
71005021	1.095
71005022	1.072
71005023	1.077
71005024	1.035
71005025	1.022
71005026	1.009
72003011	1.003
72003012	1.006
72003013	1.060
72003014	1.011
72003015	1.003
72003016	1.010
72003021	1.036
72003022	0.990
72003023	1.067
72003024	1.007
72003025	1.001
72003026	1.014
72505011	1.116
72505012	1.045
72505013	1.101
72505014	1.091
72505015	1.038
72505016	1.021
72505021	1.114
72505022	1.068
72505023	1.084
72505024	1.091
72505025	1.027
72505026	1.023
75003011	1.011
75003012	1.013
75003013	1.026
75003014	1.006
75003015	1.001
75003016	0.999
75003021	1.025
75003022	0.900
75003023	1.027
75003024	0.993
75003025	1.003
75003026	0.996
76003011	1.030
76003012	0.988
76003013	1.056
76003014	1.022
76003015	1.002
76003016	1.021
76003021	1.023
76003022	1.028
76003023	1.020
76003024	1.007
76003025	1.000
76003026	1.021
78003011	1.003
78003012	0.985
78003013	1.023
78003014	1.025
78003015	1.008

ATTACHMENT 3.—ADJUSTMENT FACTORS BY POST STRATUM [1]—Continued

Stratum code	Factor
78003016	1.001
78003021	1.041
78003022	0.967
78003023	1.016
78003024	1.016
78003025	0.997
78003026	0.961
79003011	1.013
79003012	0.995
79003013	1.041
79003014	0.999
79003015	0.993
79003016	1.001
79003021	1.010
79003022	1.000
79003023	1.015
79003024	0.994
79003025	1.001
79003026	0.997
79995011	1.077
79995012	1.007
79995013	1.082
79995014	1.065
79995015	1.033
79995016	1.021
79995021	1.069
79995022	1.026
79995023	1.725
79995024	1.048
79995025	1.027
79995026	0.969
81003111	1.034
81003112	1.035
81003113	1.123
81003114	1.066
81003115	1.041
81003116	0.990
81003121	1.041
81003122	1.061
81003123	1.060
81003124	1.019
81003125	0.977
81003126	1.013
81003211	1.031
81003212	1.021
81003213	1.035
81003214	1.020
81003215	1.002
81003216	0.966
81003221	1.031
81003222	1.020
81003223	1.045
81003224	1.003
81003225	0.990
81003226	0.960
82003011	1.017
82003012	1.017
82003013	1.071
82003014	1.014
82003015	0.978
82003016	0.975
82003021	1.030
82003022	1.021
82003023	1.064
82003024	0.998
82003025	0.967
82003026	0.962
83005011	1.066
83005012	1.023
83005013	1.107
83005014	1.063
83005015	1.027
83005016	1.005
83005021	1.058
83005022	1.055
83005023	1.077
83005024	1.025
83005025	0.973
83005026	1.004
87003011	1.022
87003012	1.008
87003013	1.076
87003014	1.023
87003015	0.988
87003016	0.986
87003021	1.026
87003022	0.984
87003023	1.022
87003024	1.000
87003025	0.971
87003026	0.965
88003011	1.021
88003012	1.032
88003013	1.059
88003014	1.028
88003015	0.990
88003016	0.993
88003021	1.036
88003022	1.028
88003023	1.036
88003024	0.996
88003025	0.971
88003026	0.995
89003011	1.050
89003012	1.027
89003013	1.077
89003014	1.036
89003015	1.031
89003016	1.003
89003021	1.046
89003022	1.049
89003023	1.041
89003024	1.024
89003025	1.017
89003026	1.014
89995011	1.110
89995012	1.076
89995013	1.123
89995014	1.070
89995015	1.039
89995016	1.062
89995021	1.106
89995022	1.064
89995023	1.105
89995024	1.057
89995025	1.076
89995026	1.049
90003111	1.043
90003112	1.076
90003113	1.098
90003114	1.094
90003115	1.004
90003116	0.963
90003121	1.047
90003122	1.056
90003123	1.089
90003124	1.014
90003125	1.015
90003126	0.977
90003211	1.017
90003212	1.043
90003213	1.022
90003214	1.011
90003215	1.000
90003216	1.012
90003221	1.037
90003222	1.041
90003223	1.019
90003224	1.031
90003225	1.000
90003226	0.995
90301111	1.142
90301112	1.075
90301113	1.115
90301114	1.124
90301115	1.127
90301116	0.971
90301121	1.105
90301122	1.020
90301123	1.069
90301124	1.008
90301125	0.971
90301126	0.960
90301211	1.160
90301212	1.056
90301213	1.126
90301214	1.116
90301215	1.094
90301216	1.013
90301221	1.132
90301222	1.070
90301223	1.106
90301224	1.074
90301225	1.042
90301226	1.054
90302111	1.083
90302112	1.055
90302113	1.159
90302114	1.095
90302115	1.068
90302116	0.992
90302121	1.055
90302122	1.080
90302123	1.106
90302124	1.035
90302125	1.030
90302126	1.007
90302211	1.052
90302212	1.000
90302213	1.042
90302214	1.004
90302215	0.990
90302216	0.990
90302221	1.029
90302222	1.016
90302223	1.094
90302224	1.025
90302225	0.971
90302226	0.995
90304111	1.047
90304112	1.053
90304113	1.147
90304114	1.090
90304115	1.074
90304116	1.046
90304121	1.069
90304122	1.060
90304123	1.068
90304124	1.074
90304125	0.961
90304126	1.057
90304211	1.076
90304212	1.052
90304213	1.071
90304214	1.065
90304215	1.028
90304216	1.029
90304221	1.070
90304222	1.072
90304223	1.079
90304224	1.026
90304225	1.033
90304226	1.029
91003111	1.035
91003112	1.045
91003113	1.112
91003114	1.073
91003115	1.020
91003116	0.986
91003121	1.000
91003122	1.033
91003123	1.045
91003124	1.020
91003125	0.947

ATTACHMENT 3.—ADJUSTMENT FACTORS BY POST STRATUM [1]—Continued

Stratum code	Factor
91003126	0.959
91003211	1.025
91003212	0.978
91003213	1.017
91003214	1.020
91003215	1.004
91003216	0.999
91003221	1.038
91003222	1.030
91003223	1.092
91003224	1.023
91003225	1.001
91003226	0.968
92003011	1.041
92003012	1.005
92003013	1.070
92003014	1.016
92003015	0.992
92003016	0.963
92003021	0.987
92003022	1.011
92003023	1.028
92003024	1.010
92003025	0.965
92003026	0.973
95003011	1.028
95003012	0.995
95003013	1.050
95003014	0.971
95003015	1.001
95003016	0.979
95003021	1.051
95003022	1.002
95003023	1.029
95003024	0.995
95003025	0.987
95003026	0.967
96003011	1.026
96003012	1.039
96003013	1.014
96003014	1.018
96003015	1.070
96003016	1.002
96003021	1.043
96003022	1.108
96003023	1.085
96003024	1.030
96003025	1.010
96003026	0.976
97001011	1.251
97001012	1.236
97001013	1.250
97001014	1.278
97001015	1.180
97001016	1.117
97001021	1.199
97001022	1.158
97001023	1.182
97001024	1.136
97001025	1.111
97001026	1.112
97002011	1.092
97002012	1.084
97002013	1.088
97002014	1.085
97002015	1.048
97002016	1.014
97002021	1.066
97002022	1.071
97002023	1.079
97002024	1.052
97002025	1.061
97002026	1.018
97004011	1.026
97004012	0.992
97004013	1.048
97004014	1.009
97004015	0.993
97004016	0.963
97004021	1.063
97004022	0.990
97004023	1.053
97004024	1.011
97004025	0.932
97004026	0.974
98003011	1.044
98003012	1.027
98003013	1.053
98003014	1.009
98003015	0.996
98003016	0.967
98003021	1.052
98003022	1.024
98003023	1.045
98003024	1.010
98003025	0.994
98003026	1.009
98004011	0.995
98004012	1.008
98004013	1.033
98004014	1.029
98004015	0.965
98004016	0.973
98004021	0.998
98004022	1.013
98004023	1.025
98004024	1.008
98004025	0.965
98004026	0.942
99003011	1.028
99003012	1.036
99003013	1.043
99003014	1.024
99003015	1.005
99003016	0.994
99003021	1.029
99003022	1.020
99003023	1.048
99003024	1.019
99003025	1.016
99003026	0.996

[1] See Attachment 2 for description of post stratum codes.

[FR Doc. 91-17202 Filed 7-16-91; 10:20 am]

BILLING CODE 3510-EA-M

ESCAP MEETING NO. 4 - 01/12/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 4

January 12, 2000

Prepared by: Maria Urrutia and Genny Burns

The fourth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on January 12, 2000 at 10:30. The purpose of this meeting was to discuss the 1990 Census experiences regarding use of statistical methods to adjust the census. Two aspects were discussed: (1) a chronology of events, and (2) the decision process.

Persons in attendance:

William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson
Jay Waite
Bob Fay
Raj Singh
Gregg Robinson
Signe Wetrogen
Carolee Bush
Sally Obenski
Maria Urrutia
Genny Burns

I. Chronology of Events for the 1990 Census Adjustment Decision

John Thompson presented the attached chronology of events for the 1990 Census adjustment decision. The following handouts for the discussion were distributed and will be on file with these minutes.

(1) Bureau of the Census, "Assessment of Accuracy of Adjusted Versus Unadjusted 1990 Census Base for Use in Intercensal Estimates," Report of the Committee on Adjustment of Postcensal Estimates, August 7, 1992.
(2) Department of Commerce, Office of the Secretary, "Adjustment of the 1990

Census for Overcounts and Undercounts of Population and Housing; Notice of Final Decision," Federal Register, Part III, July 22, 1991.

(3) Obenski, Sally, "Summary of C.A.P.E. Technical Findings", January 11, 2000.

(4) Obenski, Sally and Fay, Robert, "An Analysis of the Consistency of the 1991 Mosbacher Guidelines to Census Bureau Standards", DRAFT, January 11, 2000.

(5) Thompson, John H., Memorandum for CAPE Committee, Addendum to August 7, 1992 CAPE Report, November 25, 1992.

(6) Thompson, John H., Chronology of Events for the 1990 Census Adjustment Decision, January 12, 2000.

In 1980, the Bureau of the Census and the Department of Commerce were sued over the decision to not use statistical methods to adjust the census. Because of problems with the Coverage Measurement Survey, the Bureau was opposed to adjustment and this was upheld in court. After the 1980 Census, the Bureau formed the Undercount Research Staff.

The Census Bureau planned a dual track approach for the 1990 Census based on conducting the best possible census while having processes in place for adjustment. A decision on which track to pursue was to be made before the census. A sample of 300,000 housing units was allotted for the Post-Enumeration (PES), 150,000 of which were targeted for research and evaluation and 150,000 for adjustment purposes. In October 1987, the Commerce Department announced that the PES would not be used for adjusting the 1990 Census. This led to multiple suits being filed which were aimed at directing the Bureau to use the PES for adjustment purposes.

In 1989, a settlement was reached in litigation with the following results:

(1) The Bureau would conduct the PES and the Secretary of Commerce would decide whether to adjust the census by July 15, 1991.

(2) The Secretary of Commerce would publish guidelines that would be followed in reaching a decision.

(3) A panel of experts, four members on each side of the litigation, would be formed to advise the Secretary of Commerce.

To support the analysis of the guidelines, the Bureau conducted various evaluations. The Census Bureau senior technical staff, the Undercount Steering Committee, reviewed the evaluation results and recommended that the 1990 Census be adjusted. Director Bryant reviewed the Census Bureau's technical decision and based on the research and on her analysis recommended adjustment. Senior Department of Commerce management recommended against adjustment. The Special Advisory Panel rendered a split decision.

The Secretary of Commerce reviewed the recommendations and decided not to adjust and published this decision in the Federal Register Notice on July 22, 1991.

The Secretary of Commerce directed the Census Bureau to review the results of the PES to determine if these could be used to adjust the post-censal estimates. The Census Bureau formed the Committee on Adjustment of Postcensal Estimates (C.A.P.E.) to direct these efforts. The post-censal estimation adjustment decision was delegated to the Director of the Census Bureau. As part of the C.A.P.E. review, a computer error was found and corrected. The C.A.P.E. issued a report indicating that adjustment would make distribution for states better but could not find any differences between the adjusted and unadjusted data for entities with less than 100,000 population. Since adjustment was not demonstrated to improve coverage for all areas, Director Bryant decided not to adjust post-censal estimates on the basis of the PES. She did decide that the Federal Statistical System could use adjusted numbers for survey controls.

II. Mosbacher Guidelines

The handout, "An Analysis of the Consistency of the 1991 Mosbacher Guidelines to Census Bureau Standards," was discussed. The details of the decision made by Mosbacher are discussed in this handout. Following are the key points.

(1) The Secretary's decision was based on criteria that required the adjustment be shown to be better at all levels used. The effect of this principle was that the unadjusted census estimates were assumed to be better a priori.

(2) It was noted that the Secretary's decision rested solely on the concept of distributive accuracy. Also, it was noted that the wide variety of census data uses necessitates that both numeric and distributive accuracy are important to consider.

III. Next Meeting

The next meeting will be held on Wednesday, February 23, 2000. The agenda topics are a summary of the February 2-3 National Academy of Science (NAS) discussion on poststratification and actions the Bureau will take as a result of this discussion.

ESCAP Committee

Kenneth Prewitt
William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson, Chair
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

cc:

Teresa Angueira
Ed Gore
Ed Pike
Catherine Miller
Fay Nash
Miguel Perez
Maria Urrutia
Genny Burns
Carolee Bush
Donna Kostanich
Raj Singh
David Whitford
Gregg Robinson
Signe Wetrogen
Magdalena Ramos
Sally Obenski

ESCAP MEETING NO. 5 - 02/23/00

AGENDA

Kathleen P Zveare
02/23/2000 07:03 AM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Elizabeth Centrella/DSSD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC
cc:
Subject: Agenda for Today's ESCAP Meeting

* * * R E M I N D E R * * * *

The agenda for today's ESCAP meeting which is scheduled from 10:30-12 in Rm. 2412/3:

The summary of February 2-3 NAS discussion on post-stratification and actions the Census Bureau will take as a result of this discussion.

ESCAP MEETING NO. 5 - 02/23/00

HANDOUTS

Issues Raised at the NAS Panel Meeting, 2-3 Feb. 2000 and Census Bureau Actions

March 30, 2000

Issue: **Census Adjustment Objective/Numeric vs. Distributive Accuracy**

1. Need for clarification on the overall objective of adjustment. Is the objective to improve overall numeric and distributive accuracy or to improve demographic distributive accuracy while not adversely affecting geographic distributive accuracy?
2. Need for a clear explanation on how numeric vs distributive accuracy affect decennial planning and evaluation.
3. Need for documentation on the effects of adjustment on small areas, especially blocks.

Actions: The first two issues are being addressed in large part in the bureau's A.C.E. feasibility document being prepared by the Associate Director of Decennial Census.[1] Additionally, two analyses supporting the feasibility paper examine distributive accuracy. [2] An assessment of the Mosbacher 1991 decision criteria examines the relative merits of numeric v distributive accuracy, and an assessment of small area accuracy discusses key aspects of assessing distributive accuracy in both the census and the PES. The last issue regarding small area accuracy should be resolved upon completion of a research project on block-level accuracy conducted by Dr. Bruce Spencer and Ms. Joan Hill. [3]

Issue: **The Adjustment Decision Process/Performance Indicators**

1. Unanimous recommendation for a public discussion of the evaluation process that the Census Bureau will follow in determining whether to release adjusted redistricting data.
2. The need to identify what performance indicators/data will be available to inform decision-makers prior to April 1, 2001.

Action: The bureau will document the decision process, including identifying those data that will inform decision-makers as to whether to release adjusted data. The plan will be available for the panel's review by September 2000.[4]

Issue: **Heterogeneity**

1. An assertion was made that nothing has changed since 1990 that would affect heterogeneity. Additionally, a request was made for a document listing the changes made since 1990 that address heterogeneity issues, including a brief discussion of the expected effect of each change. This is important because with LUCA and Be Counted

and other changes, the census could be more heterogeneous than in 1990, something the A.C.E. cannot control.[5]

2. Document the strengths and weaknesses of models and assumptions used in the A.C.E.[6]

Action: Dr. Howard Hogan's staff is developing documentation to address these concerns.

Issue: Evaluations

1. The need for a summary listing of the planned A.C.E. evaluations and how they will be used to estimate total error.

Action: Such a listing can be derived from draft study plans involving the total error model drafted by PRED. The study plans should be close to or finalized by the fall panel meeting. The plan for the total error model is in progress.[7]

Issue: Poststratification

1. The need for an outline of the decision process. This would address the concern expressed that the race combination decisions seem ad hoc.
2. Choose labels carefully for poststrata.

Actions: Dr. Bob Fay will be assisting DSSD staff in documenting the decision process. Dr. Hogan will be including a rationale for the bureau plan for combining race groups.[8]

Issue: Movers

1. The need to have a brief document outlining why the bureau chose PES-C--including the weaknesses of this approach and why they are acceptable.
2. The need to look at the assumption that outmovers = inmovers numbers, that is, the need to ensure the consistency of the P- and E-samples. Therefore, the bureau needs additional information on:
    - The movement of college students from dormitories to housing units.
    - Internal migration, e.g., persons moving from FL to NY from Census Day to A.C.E. interview day. This is related mainly to the regional variable but also involves the mail, MSA/TEA, and even tenure variables.
    - Number of outmovers for which the bureau can get matchable materials.
    - Net migration among poststrata for the April to June or July timeframe.
    - Match rates for movers, before and after imputation.

Actions: Dr. Hogan will prepare a document discussing the strengths and weaknesses of PES B versus PES C and why we selected PES C. Additionally, his staff is assessing the P- and E-sample consistency issue and will include their findings in the analysis.[9]

Issue: Missing data

1. The need to approximate what would have happened in 1990 if the bureau had used the ratio estimator versus the logistic regression model.
2. Provide more details about imputation cell estimation.
3. Provide more details about the characteristic imputation
4. Noninterview (NI) adjustment to whole household (HH) noninterviews need some scrutiny (probably because not enough details in background materials)

Actions: The bureau will not approximate what would have happened in 1990 had the ratio estimator been used because there was so little missing data that the difference would have been minimial. As for the other issues, the detailed specifications in progress should address these and any other lingering concerns of panel members and invited guests. [10]

Issue: CAPI by Telephone Interviews

1. The need to address the concern that early interviews (especially by the phone) may have different expected values for missed persons and different accuracy for mover reports.
2. Is either operational or model independence being violated?
3. Is there anything to the concern that telephone CAPI will suppress reporting of children?

Actions: Dr. Hogan and his staff will assess the need to examine these issues further.[11]

Issue: Definitions/Clarification

1. Dr. Norwood requested an explanation of the differences among demographic analysis, population estimates and population projections.

Action: Dr. John Long will prepare this explanation.[12]

Issue: **The PES "Error Chart"**

1. The need to understand the numbers presented and to refute the arguments.

Action: Bob Fay and Sally Obenski are preparing a document that explains the numbers and responds to the underlying issues.[13]

End Notes

1. Thompson, John, "The Accuracy and Coverage Measurement Evaluation Survey: A Statement on the Feasibility of Increasing Accuracy Through Statistical Methods," Draft March 31, 2000.

2. Obenski, Sally, " An Analysis of the Consistency of the 1991 Mosbacher Guidelines to Census Bureau Standards," Draft February 23, 2000. Obenski, Sally, "Analysis of C.A.P.E. Findings for Small Geographic Areas," Draft February 24, 2000.

3. Spencer, Bruce and Hill, Joan, "Accuracy of Block-Level Estimates of Population," Draft XXXX.

4. TBD

5. TBD

6. TBD

7. Spencer, Bruce, "Components of Error Needed for the Total Error Model," Draft February 13, 2000.

8. TBD

9. TBD

10. Cantwell, Pat and Ikeda, Michael, ...Specifications for Missing Data Model, XXX.

11. TBD

12. TBD

13. Fay, Bob and Obenski, Sally, "An Assessment of Wachater and Freedman's PES Statistics and Issues," Draft March XX, 2000.

February 2-3 NAS Panel on Dual System Estimation: Bureau Summary of Topics, Discussion, and Closing Statements

April 3, 2000

Summary Objective and Scope

The purpose of this summary is to document the key issues that were raised and discussed during the panel's February meeting on the Accuracy and Coverage Evaluation Survey's (A.C.E.) underlying methodology, Dual System Estimation (DSE). This summary is an internally-generated analysis that identifies issues raised by the panel and invited guests to assist decennial managers in focusing discussion and prioritizing action plans. Neither the panel members nor the invited guests has seen or reviewed this document. The information in the document is not an official representation of participants' positions. It is the Census Bureau's best recollection of the discussions ensuing over the two days and will only be used for internal planning. Further, this document is not intended to be a detailed transcript with every comment attributed to a panel member or invited guest. An official document will be prepared and distributed by the panel. Consequently, other than a few of the principals, such as the panel Chair, names have not been used.

The summary reflects an analysis by Census Bureau officials of the discussion over the two-day period that has been reorganized into topics. It begins with opening statements, provides an overview of the A.C.E. design, describes key topics, associated issues, and perception on consensus, and concludes with invited guests' closing remarks.

Background

The Panel to Review the 2000 Census was convened by the Committee on National Statistics, National Research Council, in the fall of 1998, at the request of the U.S. Census Bureau. The panel, which is chaired by Dr. Janet Norwood, former commissioner of the U.S. Bureau of Labor Statistics, is charged to review the methods, procedures, and results of the 2000 census.

The panel will review features of the census that affect the completeness and quality of the data, such as the Master Address File (MAF), follow-up for nonresponse, proxy responses, race and ethnicity classifications, and other areas. The panel will also review the statistical methods, operations, and results of the planned A.C.E. and DSE methods that the Census Bureau intends to use to evaluate the coverage of the census and to produce adjusted counts in the spring of 2001. The primary focus of the panel to date has been on the A.C.E.'s design and methodologies.

To assist the panel in its evaluation, several meetings were planned to address critical aspects of the A.C.E. The panel met in October 1999 to discuss the sampling and estimation methodology of the A.C.E. In February 2000 the panel met to discuss statistical and operational issues regarding DSE. In the fall of 2000 the panel plans to meet to review the process and performance indicators that Census Bureau officials will use to decide whether to release adjusted redistricting data in the spring of 2001.

Opening Statements

Opening statements were made by the panel chair, Dr. Janet Norwood, the Director of the Census Bureau, Dr. Ken Prewitt, and the Associate Director for Decennial Census, Mr. John Thompson. (Mr. Thompson's opening statements are summarized in the numeric versus distributive accuracy section.) Dr. Norwood stated that the purpose of the meeting was to obtain as much information as possible about DSE and the A.C.E. design, *not* to revisit the 1990 adjustment issue. She further stated that the meeting was not to be about politics, that differences of opinion were to be expected, and that the panel had no official position on adjustment at this time.

Dr. Prewitt opened by telling the panel and invited guests that the Census 2000 was currently on schedule. He then made several statements about adjustment and accuracy. He stated that criticisms of adjustment can be grouped into (1) accusations that the Census Bureau is pursuing partisan politics; (2) concerns over operational feasibility; (3) concerns over public acceptance; and (4) differences in opinion about improvements realized from adjustment. He stated that the first concern is misplaced but is interested in examining the others. Consequently, Dr. Prewitt asked for a precise framing of the issues by the panel and other contributors, particularly if during that process verifiable facts can be identified. He then stated that the Census Bureau's position on distributive versus numeric accuracy was to favor numeric, as it was difficult to maximize both.

Dr. Prewitt concluded by asking two questions: (1) How would you design a census to achieve distributive accuracy? (2) What are the facts about congressional seats shifting as a result of the computer programming error discovered in 1992?

A.C.E. Design Overview

Dr. Howard Hogan, Chief of the Decennial Statistical Studies Division, was the Census Bureau presenter of the A.C.E. design. He first provided an introduction to the underlying methodology, DSE, used in the A.C.E. He included the criteria that defined an application being "in" the census: (1) appropriateness of enumeration; (2) uniqueness; (3) completeness; and (4) geographic correctness. Dr. Hogan then outlined changes and improvements in the A.C.E. over the 1990 Post Enumeration Survey (PES).

Improvements included:

- To minimize 1990 concerns about block size differences, the Census Bureau used the most recent address listing during sample reduction for the A.C.E.;
- To obviate 1990 variance concerns about small blocks, in 2000, the Census Bureau will reduce the small block universe whenever possible by using a clustering algorithm that will group smaller blocks with larger block clusters; and,
- To reduce transcription and keying errors, the Census Bureau will use Computer-Assisted Person Interviewing (CAPI) for the A.C.E. person interview. A relatively small percentage of the overall sample size will be interviewed via CAPI by telephone after their census questionnaires have been received. CAPI by telephone will facilitate training for the person interviewing that starts in July right after Nonresponse Follow-up with the Nonresponse Conversion Operation going into early fall.

Other changes included:

- Increasing the sample size and sampling probabilities should reduce variance levels from 1990 and improve small block sampling;
- Group Quarters (e.g., college dormitories) will not be included in the A.C.E.;
- Two variables will be added to the 1992 poststratification design (i.e., mail return and Metropolitan Statistical Area/Type of Enumeration Area (MSA/TEA));
- The approach for handling people who have either moved out of or into the A.C.E. sample (i.e., movers) has changed from PES B to PES C;
- Instead of searching the areas surrounding the entire A.C.E. sample for matches, the search area will be limited to a targeted extended search;
- There will be overlap between the A.C.E. and Nonresponse Follow-up during the CAPI interviewing by telephone; and,
- In 2000, because of the increase in sample size, the Census Bureau will not use the statistical method of smoothing to offset the effects of high variability.

Numeric versus Distributive Accuracy

The relative merits of numeric versus distributive accuracy was a pivotal topic in 1991 discussions about adjustment. Improving numeric accuracy refers to getting the total population as close to "truth" as possible. Improving distributive accuracy refers to getting the allocation of the population to states or other geographic units as close to "truth" as possible. Both are important to uses of census data.

Issue

While improving both aspects of accuracy is important, it may not be always possible, even at larger geographic levels. Complicating the issue is that improving distributive accuracy can apply to both demographic groups as well as to geographic units. Consequently, it is important to clarify the overall objective of adjustment, including how numeric versus distributive accuracy affects decennial planning and evaluation.

Bureau Statement

Mr. Thompson elaborated on numeric versus distributive accuracy. He explained that in planning for a census or any other survey, the Census Bureau looks for design procedures to improve numeric accuracy. If executed perfectly, such procedures will improve distributive accuracy. However, the bureau does not *plan* a design based on improvements to distributive accuracy; again, it focuses on increasing numeric accuracy. In 1990, PES evaluations almost exclusively focused on distributive accuracy while evaluations of other coverage improvement programs focused on numeric. In 2000, evaluations will focus on both aspects of accuracy for all coverage improvement programs, including the A.C.E.

Discussion

Several participants stated the importance of distributive accuracy as a goal--that is, getting the proportion of the population or share closer to "truth." In response, Dr. Hogan asked how one would go about planning for improved distributive accuracy. The Deputy Director, Mr. William Barron, asked whether the Census Bureau should walk away from numeric improvements in the American Indian count, for example, if distributive accuracy were not improved. Further, one participant pointed out that the formulae for the DSE are for counts and were not designed to improve distributive accuracy. In response, another participant stated that formulae are not determinative and that to evaluate the methodology, one must work through the counts to get to the shares and the uses of census data. Likewise, one participant concluded that everything is about shares, and another stated that operational decisions should focus on numeric accuracy, but whether to adjust should be a distributive accuracy question.

Another area of discussion was whether the appropriate share metric should be geographical or demographic. Several participants pointed out that the emphasis on counts for the undercounted population groups is really a sense of shares and is somewhat of a hybrid. From that perspective, the Census Bureau can use distributive accuracy as a planning tool. Because of increasing tension between multiple criteria, e.g., accuracy of groups v geographic areas, one participant suggested one objective could be to reduce the differential undercount without adversely affecting the rest of the census. This prompted two additional comments. One participant stated that if shares are the ultimate goal, then the Census Bureau should measure a given demographic group across all geographic areas. Another asked which vector of errors the Census Bureau preferred--demographic groups or geographical areas? He added that one vexing problem would be how to factor shares into performance indicators used to make the adjustment decision.

Consensus
Participants focused mostly on the importance of distributive accuracy but seemed to recognize the complexity of whether the objective is to improve the distributive accuracy of demographic groups or geographic units or both. Other than suggesting that the Census Bureau could plan for improved distributive accuracy by focusing on groups, no one addressed Census Bureau concerns about how one can plan for improved distributive accuracy.

Decision Criteria/Performance Indicators

In 1990, the Census Bureau conducted and assessed 21 evaluations of the PES and 11 evaluations involving demographic analysis prior to the July 1991 adjustment decision by the then Secretary of Commerce Robert Mosbacher. Synthesizing the volumes of data involved using a total error model and loss functions to determine whether the error in the census was greater than the error contained in the PES due to sampling and nonsampling (e.g., matching, recall) error. Such a comprehensive assessment cannot be conducted prior to the April 1, 2001 adjustment decision for redistricting data.

Issue
The Census Bureau will not have completed a full evaluation of the A.C.E. until well after the adjustment decision. Consequently, only limited data will be available prior to the adjustment. The Census Bureau needs to determine the degree to which it needs a formal decision process, supported by performance data from the census and the A.C.E.

The Panel Chair Statement
Dr. Norwood reiterated that the purpose of the panel is not to decide whether numbers should or should not be adjusted but to evaluate what the Census Bureau has done. She then posed the following question: What should be looked at to evaluate the accuracy of the census with and without adjustment?

Discussion
Most participants seemed to acknowledge that repeating the 1991 Mosbacher-type assessment before April 1 was not feasible. However, many did assert the need for some performance data to inform an adjustment decision, while recognizing that some subjectivity is always necessary. Dr. Norwood suggested conducting some evaluations earlier than planned. One participant stated that the Census Bureau should provide (1) a public list of things to do before the numbers are released on April 1, (2) a catalogue of assumptions and error-sources in the A.C.E. (but focus on gross errors), and (3) a list of the planned A.C.E. evaluations and how they will be used to estimate total error. Many participants provided possible performance indicators and evaluation methodologies. The suggested performance indicators included:

- Demographic sex ratios from the census;
- Address list matching from the A.C.E. housing study as an early indicator for the MAF quality;
- Census indicators, including mail response rates, geocoding errors, erroneous enumerations, duplicates, and last resort cases;
- Intercensal estimates; and
- A.C.E. indicators, including interview rates, noninterview rates, quality control measures, estimates/variance on estimates, numbers of movers, movers across poststrata lines, match rates, missing data rates, and blocks in which matches exceeded enumerations (influential blocks).

As for evaluation methodologies, one participant pointed out that the key was to understand the relation between indicators and accuracy. However, another responded by stating how difficult it will be to synthesize the errors for decision-making. Several suggested that although there was no "right" scientific approach, it was important to conduct sensitivity analyses to better understand indicators and error. Using loss functions was suggested as a means to evaluate the consequences of using adjusted versus unadjusted data at differing geographic and error levels. One participant provided the following guidance: Obtain the best possible numbers, be perceived to have done a reasonable job, and bullet-proof the Census Bureau against the unreasonable because their minds cannot be changed.

A few participants expressed a global concern that high-levels of all types of measurement error in the 1991 PES would be repeated in the A.C.E. However, others pointed out specific concerns: (1) A potential increase in erroneous enumerations due to multiple response options (especially Be Counted) and Complete Count Committees focused on making numbers bigger; and (2) A decline in data quality due to the use of Optical Character Reading (OCR) in data capture.

Consensus
Consensus was reached that the Census Bureau cannot repeat the 1991 Mosbacher-type assessment prior to the decision to adjust, but should have a documented decision process with performance indicators made publicly available. Opinions differed as to the degree and effect of nonsampling error on the 1991 PES. Near consensus was reached on the need for careful controls for Be Counted Forms (i.e., forms that do not have census identification numbers), but no consensus was reached on whether the planned use of OCR would lead to a decline in data quality.

Heterogeneity

Two concerns about DSE are heterogeneity and correlation bias. Heterogeneity occurs when there is a failure of the so-called synthetic or homogeneity assumption in producing the adjusted census counts. As the first step, the population is divided into categories or poststrata defined by a number of variables, such as age and sex, in a set of geographic areas. An adjustment factor is estimated for each poststratum. The synthetic estimate is formed by applying the adjustment factors to the corresponding counts by poststrata at the block level and aggregating the results to higher levels of geography. The synthetic assumption assumes that the probability of being missed in the census is constant for each person within a poststratum.

Residual heterogeneity, that is, heterogeneity in census inclusion probabilities not explained by the poststratification used to estimate census coverage, has two undesirable consequences. At the poststratum level and above, the consequence is correlation bias; that is, missing the same people that the census missed. For "small areas" below the poststratum level, the concern is that geographic variation in inclusion probabilities violates the synthetic assumption, leading to biased estimates for small areas. Critics of sampling often point to local heterogeneity as being an inherent problem with DSE that demonstrates why the census counts should not be adjusted. For example, the Census Monitoring Board recently published a report on how statistical adjustment fails to eliminate local undercounts.[1]

Issue

An assertion was made that nothing has changed since 1990 although a number of changes that affect heterogeneity were discussed. There needs to be determination of the degree that general documentation is needed on planned models and assumptions and specific documentation is needed on how 1991 heterogeneity concerns are being addressed in the A.C.E. design.

Bureau Statement

Dr. Hogan explained that the initial design did not share undercount data across state lines. Undercounts for states were estimated directly from state-based samples, mitigating heterogeneity concerns. After the Supreme Court decision disallowing sampling to be used for apportionment purposes, the Census Bureau reduced the sample size and changed the design by allowing the sharing of undercount data among states. Consequently, Dr. Hogan launched an intensive research effort to design a poststratification model that would reduce heterogeneity.

Discussion

A presentation was given by Mr. Charles Jones, a Census Monitoring Board staff member, on the results of the Board's Report to Congress. [2] The presentation focused on the effects of adjustment on 1990 block-level accuracy. The statement was made that DSE methodology does not correct small area problems because of heterogeneity. The study used directly calculated block-level DSEs to test the assertion. However, it was pointed out by other participants who had assessed the study that the block-level data used by the Board did not have the associated weights and were not the numbers used in 1990 to form the synthetic estimates. Further, one participant dispelled the notion that the direct DSE was "truth," by giving Ft. Polk as an example. Although Ft. Polk has a large population of

---

[1] U.S. Census Monitoring Board, "Unkept Promise: Statistical Adjustment Fails to Eliminate Local Undercounts, as Revealed by Evaluation of Severely Undercounted Blocks From the 1990 Census Plan," Report to Congress, September 30, 1999.

[2] Ibid.

multiple minority groups, instead of indicating a high undercount as expected, the direct DSE indicated low undercount. Additionally, the participant stated that the Board's study focused on the "tails" of the DSE distribution (i.e., on the extremes) and, if one focused on the bulk of the distribution, it made the case that the DSEs were, in fact, behaving as expected. That is, although the effect was small, adjustment did move the block-level counts generally in the right direction. Moreover, the block-level data are aggregated to higher levels where definitive improvements can be demonstrated.

Several participants generally seemed to like the mail response variable in the poststratification scheme and believed that it would reduce heterogeneity. An assertion was made that little had changed since 1990 that would affect heterogeneity. One participant asked whether the Census Bureau was planning to document how heterogeneity and correlation bias are or are not dealt with in the A.C.E. He suggested that perhaps the research community should spearhead a major effort of this kind. Generally, however, statements indicated that little difference exists between unadjusted and adjusted data at the block-level--both have errors. Focus should not be on blocks but on how blocks are aggregated into tabulations for census uses.

Consensus

No consensus was reached on how serious the effects of heterogeneity are on DSE accuracy.

Poststratification

The 1991 PES design included 1,392 different poststrata. Poststratification is the dividing up of the population into groups with similar capture probabilities. They were formed according to pre-identified variables, such as age, sex, race, tenure, and other variables thought to be associated with differing capture probabilities. A person could be only one poststratum grouping. One of the complexities for 2000 is that the ability for people to select more than one race increases the possible outcomes by almost tenfold.

Because the sample size was relatively small (about 160,000) and the number of poststrata large (1,392), some of the poststrata had high variances that led to the use of a statistical model called smoothing. In part, because of the complexity of smoothing, the Census Bureau eliminated its need by reducing the number of poststrata to 357 during its 1992 analysis of whether intercensal estimates should be adjusted.

Issue

Two issues emerged over poststratification. The first was the adequacy of the Census Bureau's design. However, the Census Bureau made a late change in a variable and was not able to provide detailed performance information on the selected model. A second issue was the treatment of individuals who respond to more than one category. A potential problem is that, given different data collection modes between the census and the A.C.E. (i.e., paper versus CAPI) and as the number of choices increase (e.g., selecting single race, two races, or three), misclassification errors could increase.

Bureau Statement
Dr. Hogan began the poststratification discussion by stating that the 1992 357-poststrata design was used as the baseline design for the A.C.E. The major changes included (1) using region only for non-Hispanic White owners; (2) adding a mail response variable (the real-time mail response rates); and (3) attempting to approximate urban/rural through the MSA/TEA variables. He explained that originally his staff was hoping to get urban/rural status but was told that it would not be operationally possible. Therefore, his staff had to make a late change in the plan, which will be finalized and documented in a couple of weeks. He then identified the following criteria used for selecting poststrata:

- Similar capture probabilities;
- Similar net undercount;
- Permit detection of differences among geographic areas;
- Poststratum cells should > 100 A.C.E. sample cases;
- Operationally feasible to implement in time;
- Minimize classification error;
- Account for changes in the census since 1990; and
- Explainable.

Discussion
A number of questions arose on poststratification, including concerns about the late change in the urbanicity variable, the poststrata sample size, the selection criteria, and the handling of muti-racial responses. Regarding the late change in the urbanicity variable, Dr. Hogan again explained that it was due to determining that obtaining information on urban/rural was not feasible. Therefore, his staff was attempting to approximate the urban/rural variable indirectly through MSA/TEAs. He also explained the variability in sample size (i.e., enough sample to support a robust sample yet include the >100 criteria). Most strata will have large samples. The lower limit of greater than 100 was due to concerns about small, geographically disperse groups, such as American Indians not on reservations and Hawaiians and Pacific Islanders. When asked whether he had looked at combining strata to address this issue, Dr. Hogan stated that he and his staff had looked at several combining options but discarded them to keep a balance between variance and bias. As for determining which poststratification scheme is superior, Dr. Hogan explained that his staff had conducted simulations, looked at variances, and used targets to measure bias.

When asked about the multi-race issue, Dr. Hogan briefly provided some insight into how the muti-race responses will be handled. For example, he stated that American Indians on Indian Country will be coded as an American Indian regardless of having selected another race and/or Hispanic Origin. Any person marking Black and another single race group is coded Black, but any person marking Asian and White are coded White.[3]

Dr. Hogan readily admitted that there is not much empirical data to support the selected plan. However, Dr. Hogan pointed out that the importance of minimizing classification error cannot be overstated and that, as such, we were using our professional judgement in attempt to control misclassification. It was generally concluded that little knowledge exists on combining racial groups, but a general unease was expressed by several participants that the Census Bureau was taking on such a *de facto* role in deciding how to group racial responses.

Consensus

Because of the late change in the poststratification model, detailed information on how the final model was selected and performance data were not available for participants' review. Consequently, no consensus was reached on the adequacy of the model other than a general agreement that the mail response variable should reduce heterogeneity. As for the multi-race issue, although a general wariness about the Census Bureau plan to handle multi-race responses was expressed by participants, no one seemed to know what to do about it except to suggest caution when choosing poststratification labels.

Treatment of Movers--PES C

People who move present a special challenge for designing a DSE for census application for two reasons. First, people who move are more likely to be missed by the census and by the survey. Second, if a person has a different "usual residence" at the time of the survey than he did at the time of the census, one must decide where to sample him. In the 1990 PES, movers were sampled where they lived at the time of the survey interview. The Census Bureau then searched the census records at, and only at, their April 1 usual residence. This is known as "procedure B" or "PES B." For census 2000, a different procedure will be used, known as "procedure C" or "PES C." The A.C.E. will estimate the number of movers by the number of people who moved into the sample blocks between April 1 and the time of the A.C.E. interview (in-movers). In PES C, the Census Bureau will attempt to determine who lived at the interview address on April 1. If the residents have moved, then interviewers will have to obtain information on them from proxies; that is, either from the new residents or neighbors.

Issue

The move from PES B to PES C represents a significant change in sampling methodology from 1990 that results in design trade-offs rather than improvements. The trade-offs include

[3]For a detailed description of the current plan see, DSSD Census 2000 Procedures and Operations Memorandum Series #Q-21, "Accuracy and Coverage Evaluation Survey: Poststratification for Dual System Estimation," January 12, 2000.

(a) increased simplicity but may result in inconsistencies between the P- and E-samples, and
(b) easier matching but may result in more response error.

Bureau Statement

Dr. Hogan explained the chief differences and the trade-offs between PES B, used in 1990, and PES C, planned for 2000. He pointed out that for PES B, the advantage is talking directly to the person while the disadvantage is that matching is difficult. As for PES C, matching is easy with a reduction in geographic matching error. However, there are more proxy interviews and higher noninterview rates, which could affect the quality of DSE. He stated that there aren't more inmovers than outmovers because the assumption is that the overall number of inmovers equals the overall number of outmovers. PES C tries to compromise between approaches to measure the number of movers and enumeration rate for movers.

Discussion

Most of the discussion focused on the assumption that inmovers and outmovers balance and on the quality of the proxy interviews. The first issue centered on regional differences, seasonal movements, and college students. One participant pointed out that the Northeast Region, for example, doesn't have nearly the inflow of people that the South and the West do. Likewise, another asked about seasonal movements. Dr. Hogan responded to the former by stating that he cannot believe that net migration over 2-3 months is so large as to make assumption about movers problematic. He responded to the latter by stating that seasonal movement was probably not true poststratum to poststratum, and that it is the large net flows between April and July crossing poststrata that inflate or deflate measures; hence, the rates should balance. The last concern was about college students coming out of dormitories not included in the A.C.E. sample because they are Group Quarters and moving into housing units for the summer that could be in the P-sample. One participant questioned whether the P-sample was consistently defined, due to this issue with college students and whether the Census Bureau could compare the match statuses between PES A (similar to PES C) and PES B using 1990 data.

Several participants expressed concern about the proxy interviews. Dr. Hogan explained that tracing is a complex and difficult task that was attempted in the dress rehearsal but was not practical for 2000. Further, tracing proved to be ineffective in the dress rehearsal. One participant pointed out that proxy data could increase response error and hence underestimate the match rate, leading to an inflated undercount.

Consensus

General consensus was reached that these two issues--consistency of the P- and E-samples and effect of proxy data on response error--need to be closely examined and documented.

Missing Data

As in all surveys, there will be nonresponse and incomplete response at various steps. The goal of the missing data process is to improve the DSE estimates. In choosing missing data procedures, the Census Bureau chooses methods that support the underlying DSE assumptions. In 1991, the Census Bureau used a fairly sophisticated hierarchical logistic regression model. Although scrutinized by critics, two different evaluations validated the model. Moreover, missing data rates in 1991 were negligible.

Issue

The Census Bureau has moved from the 1990 logistic regression model to a simpler ratio estimator model that may not be as accurate. Further, in 2000, missing data may be more of a problem than in 1990.

Bureau Statement

Dr. Hogan stated that imputations will be based on all available information and described the three types of missing data found in the A.C.E.: (1) whole household noninterviews; (2) missing characteristics (needed for poststrata) using the hot deck methodology; and (3) missing enumerator sample. He explained that in 1990 the Census Bureau used a hierarchical logistic regression model, but in 2000 it would use a ratio estimator although the specifications are not yet finalized.[4] Howard further explained that the cell model is easier to verify and to program, which are important considerations.

Discussion

Several participants expressed concern about the move to a ratio estimator. One stated that he was not happy with the move to a unit nonresponse adjusted weight due to the limitations of the ratio estimator. He acknowledged that it was selected for simplicity but stated that there is a danger in making things too simple. He stated that a detailed explanation of the imputation model may help. One participant who had been critical of the regression model in 1990 stated that he would support the Census Bureau on using the ratio estimator for missing data. He stated that he doesn't like the logistic regression model because of the way it treats the bias/variance tradeoff-- at the expense of variance. Another participant pointed out that there is bias in the cell model too, but no sense of variance. He questioned whether we could develop a hybrid, that is use all interactions for some variables and only main effects for others. A final participant commented that the 1990 evaluations showed that the logistic regression model was surprisingly effective. The participant asked if the Census Bureau knew the effect on 1990 data if the cell model were used in lieu of the regression model and wondered what is really being gained or lost.

[4]According to the Master Activity Schedule, the detailed specifications will be completed by April 17, 2000.

Consensus
No consensus was reached on the effect of changing models. Generally, participants stated that they would hold judgement until detailed specifications on the missing data model were available.

Telephone CAPI and Interview Duration

In 1991, the Census Bureau did not begin the PES interviewing until the end of Nonresponse Follow-up. In 2000, primarily to ease schedule concerns, the Census Bureau will begin CAPI interviewing by telephone in block clusters that have been enumerated. However, it will overlap with Nonresponse Follow-up in other areas. Additionally, the full data-gathering phase for A.C.E. will extend from May of 2000 well into the fall.

Issue
The length of time between the A.C.E. interviews and Census Day and possible differences between the early telephone CAPI respondents and later respondents could cause an increase in response error and heterogeneity. Additionally, because of its overlap with Nonresponse Follow-up, the use of telephone CAPI could violate independence assumptions. Finally, telephone CAPI could exacerbate the undercount of children, who were a large percentage of the 1990 undercount. Some empirical studies suggest particular difficulty in identifying children as household members through telephone interviewing.

Bureau Statement
Dr. Hogan described the sequence of operations in the A.C.E. First, CAPI telephone interviewing would occur during Nonresponse Follow-up but would include only a small universe, i.e., 10-15 percent, of the cases. Second, the CAPI person interviewing occurs using the same interviewers. Third, the Nonresponse Conversion Operation (designed to resolve nonresponses) occurs. Finally, person follow-up to resolve nonmatches occurs well into the fall. [5] He further explained that the A.C.E. uses the respondents' belief of where they should have been enumerated on Census Day, rather than trying to impose a precise definition on ambiguous cases.

Discussion
Two areas of concern were raised--(1) possible errors due to the lag between Census Day and the A.C.E. interviewing and due to possible differences between telephone CAPI responses and later responses, and (2) the effect of the use of telephone CAPI on independence and responses about children. Addressing the first area of concern, Dr. Hogan explained that the Census Bureau has no

---

[5] CAPI by telephone begins on May 8 and CAPI person interviewing begins on a flow basis as each block cluster is enumerated and ends on August 19. Nonresponse Conversion begins on July 27 and ends on September 1. Person follow-up begins on October 23 and will end on November 21.

data and would have to conduct a study within a study to see if there were differences between early and later respondents, which was not feasible for 2000. Dr. Norwood asked about data from other surveys and Dr. Hogan explained that the Census Bureau had used data from studies to design address and follow-up probes to compensate for recall error. Another participant pointed out that there was a correlation between the timing of interview and capture/response. Regardless of CAPI by telephone, when a survey gets to someone late, people are more likely to move or have a response problem. Therefore, there is an underlying correlation between characteristics of the household and quality. He asked if the Census Bureau had planned experiments on this. To this, a participant commented that people see heterogeneity everywhere and asked why this is even an issue. He pointed out that it sounds like an interesting question but has little effect on estimation.

As for the second major concern--the use of telephone CAPI--one participant asked if telephone CAPI violated the independence assumption. He elaborated by asking if telephone CAPI increases the chance of some people to be included in the P-sample. If they are treated differently, he asserted their data capture probability would be increased and that data quality would differ. Another participant responded that he believed that operational independence is not the issue. He thought that telephone CAPI may compromise model independence, may affect movers, and consequently, evaluation is needed.

As for telephone CAPI and children, one participant pointed out that in 1990, 50 percent of the undercounted were children and that many were left off of the census form and asked if the Census Bureau is concerned about the use of telephone CAPI. Another participant pointed out that the National Immunization Survey results demonstrated how hard it is to find 2-year-olds. A third participant asked what the Census Bureau was doing to increase children's response rates. Mr. Thompson responded that they had initiated a very extensive Census in Schools project. Further, there are quality checks built into the census to ensure that individuals are not left off the census form. For example, the "number of people in the housing unit" number is checked against the number of people listed on the roster.

<u>Consensus</u>

General consensus was reached that these two issues, i.e., possible response differences and the effects of telephone CAPI on the inclusion of children, were legitimate concerns that should be examined.

Closing Comments

The following information represents the Census Bureau's interpretation of key closing comments made by participants. The purpose is informational and the statements and facts presented have not been reviewed by the participants nor have they been verified by the Census Bureau. Consequently, other than the final remarks made by Dr. Prewitt, the names of the participants have been omitted.

Participant 1

- Census 2000 is similar to 1990--and the adjustment is similar. Nothing has really changed.

- To get a handle of magnitudes of correlation bias (CB) and measurement bias (MB) look at 1990 as illustrated by the following chart.

<u>1991 Mosbacher</u>

The adjustment	=	5.3M	
MB	=	3.6M	1.8M (computer error)
Corrected Adj	=	1.7M	
CB	=	3.0M	
DA	=	4.7M	

MB numbers CAPE	=	3.0M
Breiman	=	4.2M
Intermediate	=	3.6M

- Given these levels of error, I question whether any improvement could be made.

- If the Census Bureau had adjusted in 1991, every single state would have gotten an upward adjustment. If two-thirds of the "undercounted" people are really the result of measurement bias, adding people seems insecure. Even with the 1.7 million undercounted, the Census Bureau doesn't know what states they should be allocated to--only what poststrata they are in.

- Even an increase in numeric accuracy is not obvious.

- As for distributive accuracy, in terms of shares, California, Texas, and Florida got upward adjustments while Pennsylvania, Ohio, and some other states got downward adjustments. Given the level of error in the adjusted numbers, it would have been just as plausible if the signs had been reversed.

- How would the Census Bureau defend the production adjustment versus the alternative adjustment cited above? It is difficult when the PES was dominated by measurement bias and heterogeneity. The same problems remain for census 2000.

- In answer to Dr. Prewitt's questions made in his opening statement:

  (1) The outcome in 1990 with regard to improvements in adjustment to distributive accuracy was that distributive accuracy had a way to go. As to how to design a census to achieve

distributive accuracy? I will talk about that for 2010.
(2) As for moving congressional seats, the 1991 adjustment of 5.3M would have shifted 2 seats; in 1992, when the Census Bureau changed the poststratification design and corrected errors, only 1 seat shifted.

Participant 2

- I recommend explaining DSE using Rick Griffin's paper (page 3)[6] rather than the 2x2 table used today.

- Those people with close to zero probability of being counted either in the census or in the A.C.E., the unreachables, are always going to be a problem. We need to focus on problems that we can control.

- In 1990 the Census Bureau attempted to address all sources of error in the total error model. In 2000, addressing total error doesn't seem quite so integrated into the census process. However, total error analysis is needed for understanding error components and the total error model even more for 2000.

Participant 3

- The Census Bureau should make evaluation data available at the lowest level.

- The Census Bureau should conduct loss function analysis using a total error model.

- The Panel should assess the A.C.E. evaluations looking especially at movers and missing data.

Participant 4

- Based on the performance of other similar systems, I am still concerned about the effect of OCR on quality. No human will be randomly looking at data quality.

- My deepest fear is that changes in the 2000 design will negatively impact census accuracy: multiple modes, OCR, new differential coverage programs such as the Complete Count Committees.

---

[6]DSSD Census 2000 Procedures and Operations Memorandum Series #Q-20, "Accuracy and Coverage Evaluation Survey: Dual System Estimation," page 3, January 12, 2000.

- If census accuracy is compromised, this will favor adjustment. However, we will need quite a bit of information to determine if an improvement is made with adjustment. Therefore, the Census Bureau must have a set of evaluations completed prior to adjustment. Otherwise, there is no way to determine which is better.

Participant 5

- The Census Bureau needs early evaluations of census counts. I am still concerned about multiple modes increasing the probability of duplications, i.e, erroneous enumerations. It is important to look at the census and make a determination on what the Census Bureau expects the adjustment to do.

- I recognize that the Census Bureau cannot do all the evaluations before the adjustment decision; it would be useful to do select ones to get some sense of the quality of the A.C.E.

Participant 6

- The Census Bureau needs to define things objectively. For example, get away from the definition that by correct residence it is "belief of residency." The Census Bureau needs errors in the measurement of the estimate objectively defined.

- As for PES C, two errors are possible--error in the estimate of inmovers means error in estimate of outmovers.

- The Census Bureau needs evaluations in real-time to support the adjustment decision; look at a full set of indicators, including demographic analysis sex ratios, numbers, and postcensal estimates.

Participant 7

- The total error model is of the utmost importance in evaluating the A.C.E. and the DSE. As for indicators to help inform decision-making, synthesizing the indicators is the challenge.

- In the census, the Census Bureau should look at last resort statistics and the results of the Primary Selection Algorithm to get an insight into duplicates.

- In the A.C.E., look at match rates, the return rate for listed blocks, geocoding error, duplicates, and extreme blocks (influential observations).

- By looking at a set of indicators, it could give insight into the synthesizing problem.

Participant 8

- The Census Bureau should provide as much data as possible for review by the statistical community. The more data provided, the better.

- I like loss function analysis although the Census Bureau's choice of loss functions in 1991 was subjective.

- Increased publicity and the expanded Census in Schools in 2000 are positive changes. The move to multi-race and treatment of movers are new and may possibly be problematic.

Participant 9

- I am concerned with overcounts.

- I suggest that the Census Bureau (1) summarize census and adjustment methodologies; (2) identify their strengths and weaknesses; (3) determine unknowns; and (4) assess which of the unknowns are possible to resolve.

Participant 10

- Here is a situation where one should be concerned about all the errors, not just net errors. If there were 25M gross errors as described by Participant 1, even if the net is 0, adjustment would still be worthwhile.

- Also, per Participant 1's statement about changes in congressional seats going from 2 to 1 from 1991 to 1992, participant 1 left out an important piece of data. He stated that when the 1992 357-design was run without the computer error, only 1 seat shifted. However, if the 1392 design is run again without the computer error, 2 seats still move. I discuss this and other topics on my website (http://lib.stat.cmu.edu/~fienberg/WhoCounts.html).

- Block level accuracy is largely irrelevant; it is primarily a device to add up to larger geographical units.

- I was impressed by the thoroughness of the material provided by the Census Bureau, but note that there were no citations outside the Census Bureau.

- I have two major concerns about the census process: (1) Local efforts that could increase erroneous enumerations; and (2) a possible bias (i.e., less long-form data) in the long form due to people selecting unlabeled short forms through multiple modes, such as Be Counted.

- I have identified the following issues: (1) taking stock of local improvement efforts (overcount issue); (2) conducting long-form evaluations possibly using data from the American Community Survey; (3) identifying evaluation criteria--that is, a public list of things to do before the Census Bureau releases numbers on April 1, 2001, including a list of formal evaluations; and (4) evaluating the American Factfinder's confidentiality. On the latter, I am concerned about the disclosure issue for race.

Dr. Prewitt's Closing Remarks

- Social recognition of groups is a numeric count issue and an important concept.

- I am aware of the overcount issue. The Internet is probably going to be fine because Internet forms require census identification numbers. Blank forms, like Be Counted, will require internal checks.

- The decision on whether to use the DSE to adjust has to be made in real-time with limited input.

- Four potential outcomes could be considered in this decision process.
  - If the A.C.E. is good, and the census is bad--adjust.
  - If the A.C.E. is bad, and the census is good--don't adjust.
  - If both the A.C.E. and the census are good--may or may not adjust.
  - If neither the A.C.E. or the census is good--very tough decision.

Conclusion

Although the panel Chair made it clear that the purpose of the meeting was not to "relive" 1991, that proved to be difficult for two reasons. First, the 1991 PES, the DSE methodology, and the adjustment controversy were the impetus for the emergence of a plethora of complex policy and technical issues generating numerous technical papers over the decade. Second, the A.C.E.'s expected performance is premised on the 2000 census and the A.C.E. performing comparably to the 1990 census and PES. Therefore, many of the 1991 issues about adjustment remain relevant. As such, the preponderance of issues raised by the panel were comparable to ones raised during the 1991 adjustment controversy: How will the Census Bureau know whether it is prudent to adjust? Which is more important, distributive or numeric accuracy? Will heterogeneity be a problem? Are the underlying statistical models valid and robust?

In 2000, unlike 1991 where the Census Bureau was breaking new ground, the Census Bureau has the advantage of a decade of intense research and test results to draw upon. Consequently, anticipating many of the concerns raised during the meeting, the Census Bureau has been in the midst of, or will be documenting, many of the issues raised by the panel, as well as their resolution.

ESCAP MEETING NO. 5 - 02/23/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 5

February 23, 2000

Prepared by: Maria Urrutia and Genny Burns

The fifth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on February 23, 2000 at 10:30. The agenda for the meeting was a summary of the February 2-3 National Academy of Sciences (NAS) discussions on Census 2000 and actions the Bureau will take as a result of these discussions.

Persons in attendance:

William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
John Thompson
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
Susan Miskura
Tommy Wright
Raj Singh
Gregg Robinson
Signe Wetrogen
Carolee Bush
Sally Obenski
Maria Urrutia
Genny Burns

I. Overview of the Meeting with the National Academy of Sciences (NAS) Panel

John Thompson distributed a summary prepared by Sally Obenski on the meeting with the National Academy of Sciences (NAS) Expert Review Panel on Census 2000 on February 2-3, 2000. He also distributed a list prepared by Bureau staff of policy and technical issues raised at the meeting. These handouts will be included with these minutes and kept on file.

The high points of the summary and list of issues were briefly discussed. John asked that the Committee read the handouts and give comments to Sally especially on topics that may have been omitted. These comments will be incorporated into a document clarifying the issues and responses, decisions, and actions on the major concerns. John also asked Bob Fay and Sally Obenski to gather more information on distributive accuracy in undercounts and expand the summary to include this work.

A list of the topics discussed at the NAS meeting is as follows:

Numeric versus Distributive Accuracy
Performance Indicators/Adjustment Criteria
Heterogeneity
A.C.E. Design Overview
Poststratification
Treatment of Movers–PES C
CATI and Interview Duration
Missing Data

There seems to be broad consensus of all participants that the Census Bureau needs open discussions on whether to release adjusted data and the processes/criteria that will be followed during this process.

A number of issues were also discussed as described in the handout. These issues were discussed during this ESCAP meeting and the attachment describes the actions that will be taken.

II. Next Meeting

The next meeting scheduled for Wednesday, March 8, 2000 has been canceled and is currently being rescheduled.

<u>ESCAP Committee</u>

Kenneth Prewitt
William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson, Chair
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

cc:

Teresa Angueira
Bill Bell
Debbie Bolton
Genny Burns
Carolee Bush
Gerald Gates
Ed Gore
Dave Hubble
Donna Kostanich
Ellen Lee
Charlene Leggieri
Don Malec
Betsy Martin
Catherine Miller
Fay Nash
Sally Obenski
Miguel Perez
Ed Pike
Magdalena Ramos
Gregg Robinson
Raj Singh
Maria Urrutia
Signe Wetrogen
David Whitford
Henry Woltman
Tommy Wright

ESCAP MEETING NO. 6 - 03/22/00

AGENDA

Kathleen P Zveare
03/21/2000 10:40 AM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Elizabeth Centrella/DSSD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC
cc:
Subject: Agenda for March 22 ESCAP Meeting

The next ESCAP meeting will be March 22 from 10:30-12 in Rm. 2412/3. The agenda is as follows:

Overview of A.C.E. 2000 Evaluations - Ruth Ann Killion/Howard Hogan

ESCAP MEETING NO. 6 - 03/22/00

HANDOUTS

The
Census 2000
Evaluation Program

Prepared by the
Planning, Research, and Evaluation Division

February 4, 2000

CENSUS 2000 EVALUATION PROGRAM
Table of Contents

February 4, 2000

Introduction[1]

For over half a century, the Census Bureau has conducted a formal evaluation program in conjunction with each decennial census. For Census 2000, the Evaluation Program will assess the effectiveness of key operations, systems, and activities in order to *evaluate the current census* and to *facilitate planning for Census 2010* and the *American Community Survey*.

The Census 2000 Dress Rehearsal, conducted in 1998, included evaluations of questionnaire design, field operations, data processing, and estimation. Over 40 evaluation studies were used to inform the final Census 2000 design. The Census 2000 Evaluation Program more than triples this effort; about 140 evaluations are planned.

These evaluations fall into 18 broad categories covering response rates, data quality, promotion and partnerships, address list development, field operations, coverage improvement, data capture and processing systems, the Accuracy and Coverage Evaluation survey, and others. The evaluations speak to issues of quality, plausibility, feasibility, accuracy, effectiveness, and value, and will provide a comprehensive assessment of the operations and outcomes of the census. For each of these 18 categories, this document provides an "Overview" and a "What Will We Learn" section, followed by a brief description of each planned evaluation.

In addition to the evaluation studies, the Census Bureau will prepare a variety of profiles or assessments based on real-time operations. For example, a quality profile based on the address listing quality assurance results will be prepared.

The design of Census 2000 is by far the most ambitious decennial census in history, particularly in its uses of an open planning process, promotion, partnerships, new technologies, statistical methodology, and alternative methods for hard-to-count populations and areas. Yet, as our nation continues to grow in size, complexity, and the need for rapid and accurate data, all these things and more will need to be further refined and developed to meet the challenges of providing data in the 21st Century - more data needs at lower levels of geography on a more timely basis.

The Census 2000 Evaluation Program is an ambitious program to assist the Census Bureau in evaluating Census 2000 and in exploring new survey procedures in a census environment. In conjunction with the Census 2000 Testing and Experimentation Program, it will build the foundation for making early and informed decisions about the role and scope of Census 2010 in the federal statistical system and its interaction with the American Community Survey and the Community Address Update System. This work will provide critical analysis and information for Census Bureau planning and implementation decisions for Census 2010 and the American Community Survey.

---

[1] The scope of the Census 2000 Evaluation Program is not final; therefore, the information presented in this document may change.

This page intentionally left blank.

A: Response Rates and Behavior Analysis

Overview

These evaluations examine various modes for providing responses to the census. We will study the use of the telephone and Internet as response options along with their use in providing assistance to respondents. The effectiveness of mailing practices and the targeted dissemination of forms will also be assessed. These evaluations focus on respondent behavior and how that behavior impacts response rates (i.e. mailback, telephone, and Internet). Findings from these evaluations will identify methods that can be used in future censuses to improve the overall response rates.

What Will We Learn?

The findings from these evaluations will answer a number of critical questions about how quickly and completely the U.S. population responded to Census 2000. From a technical standpoint, the use of an Internet Questionnaire Assistance (IQA) module will demonstrate the utility of employing the "most current" technologies and provide insight into respondent perception of using this mode for requesting information or completing a questionnaire. Likewise, an enhanced telephone questionnaire assistance (TQA) program that is user-friendly and comprehensive will provide further insight into respondent needs and preferences. Evaluation of the TQA program also will provide insight into questions concerning contractor support.

Analyzing mail response/return rates (by form type, demographics, and geography) and mailing practices, such as tracking undeliverable questionnaires, will provide insight into improving overall response rates. Assessment of the *Be Counted Campaign* will help determine the benefits of targeting geographic areas and/or demographic groups in an effort to improve population and housing coverage. We also will examine the frequency of use of language assistance guides and questionnaires in languages other than English, along with the number of returned non-English questionnaires.

(A.1.a) Telephone Questionnaire Assistance Operational Analysis
The Census 2000 Telephone Questionnaire Assistance system was developed with contractor support to provide the following services to respondents: 1) helping them complete questionnaires, 2) providing questionnaires (English forms only) and foreign language guides upon request, and 3) conducting short form questionnaire telephone interviews when necessary. This operational evaluation assesses respondent behavior (i.e., calling patterns), the accuracy of geocoding results, and the quality of data received through the various modes.

(A.1.b) Telephone Questionnaire Assistance Customer Satisfaction Survey
This evaluation focuses on customer reaction to the Census 2000 Telephone Questionnaire Assistance program. It includes analyses in the following areas: accessibility, ease of use, overall satisfaction with the assistance, and appropriateness of the information provided.

(A.2.a) Internet Questionnaire Assistance Operational Analysis
When accessing the Census Bureau website, the Census 2000 Internet Questionnaire Assistance program provides an informational service to respondents. This is the first time such a service has been available to the public. This evaluation assesses the type of service requested, the total number of visits to the site, the time (date, day of week) distributions of these visits, and the topics/pages most frequently searched. Note that Internet Web hits are a poor measure of traffic volume, but in most cases they are the only measure available. They can be used as a relative measure of one page's hits relative to another page's hits, or one server's hits relative to another server's.

(A.2.b) Internet Data Collection Operational Analysis
For Census 2000, respondents have the opportunity to complete the short form questionnaire on the Internet. This is the first time a decennial census has used this data collection mode. Since there is no background data on what might be expected in terms of frequency of use and completeness of the data, a general evaluation of the Internet data collection mode is planned.

(A.2.c) Internet Website and Questionnaire Customer Satisfaction Survey
Customer satisfaction surveys will be used to examine the effectiveness of both the Internet Questionnaire Assistance and the Internet Data Collection programs.

(A.3) Be Counted Campaign
The Be Counted Campaign makes blank questionnaires available at convenient locations for persons who believe they may have been left out of Census 2000. This evaluation will examine person and housing unit coverage gains from the campaign along with the characteristics of those enumerated on Be Counted forms. This evaluation also will assess the impact on the Master Address File through documentation of housing unit adds resulting from this program, and it will evaluate our ability to geocode and process Be Counted forms.

(A.4) Language Program - Use of Non-English Questionnaires and Guides
This study will document how many housing units were mailed the advance letter about requesting a non-English questionnaire, by state and type of enumeration area (e.g., mailout/mailback, update/leave, etc.); how many non-English forms were requested, completed, and checked in; and the frequency of requests for non-English short and long forms. This study also will document the number of language assistance guides requested through Telephone Questionnaire Assistance, Questionnaire Assistance Centers, and the Internet, along with an analysis of which languages were most often requested, whether the requests were clustered geographically, and how many requests for a language assistance guide resulted in a mail returned form.

(A.5) Response Process for Selected Language Groups
This evaluation will provide insight into how Spanish, Vietnamese, and Russian speaking households coped with the census questionnaire in Census 2000. Specifically, we will look at how these non-English speaking long form households were enumerated. We will assess their use of language guides, Questionnaire Assistance Centers, Telephone Questionnaire Assistance, and their experience with the English form.

(A.6.a) U.S. Postal Service Undeliverable Rates for Census 2000 Mailout Questionnaires
For Census 2000, the questionnaire mailout/mailback system provides the primary means of enumeration. This type of enumeration is conducted mainly in urban and suburban areas, but also in some rural areas that contain city-style address (house number/street name) systems. This evaluation examines the rates at which housing units were classified by the U.S. Postal Service as "undeliverable as addressed" (UAA) for varying levels of geography; the occupancy status of those housing units; demographic characteristics for housing units that were deemed undeliverable but had a final status of occupied; the effect that undeliverable questionnaires had on nonresponse rates; and the check-in pattern of UAA questionnaires according to date of receipt.

(A.6.b) Detailed Reasons for U.S. Postal Service Undeliverability of Census 2000 Mailout Questionnaires
This evaluation further examines the issue of the undeliverability of census mailout questionnaires. After the U.S. Postal Service determines that mail pieces are "undeliverable as addressed" (UAA), the Census Bureau will attempt to deliver these cases at the Local Census Office level. This evaluation assesses the quantity of questionnaires designated as UAA and the distribution of the UAA questionnaires according to reason for undeliverability.

(A.7.a) Census 2000 Mailback Response Rates
Housing units in mailout/mailback and update/leave enumeration areas are asked to return questionnaires in postage paid envelopes. Those questionnaires are received and checked in at Data Capture Centers. This evaluation examines mail response rates at varying levels of geography and quantifies information about incoming questionnaires according to form type and timing with respect to critical operational dates.

(A.7.b) Census 2000 Mail Return Rates

Housing units in mailout/mailback and update/leave enumeration areas are asked to return questionnaires in postage paid envelopes, and once all followup operations are complete, those housing units are assigned a final status. Only the housing units that were assigned to receive an update/leave or mailout/mailback questionnaire and had a final status of *occupied* are factored into the mail *return* rates. Data on mail return rates provides more accurate measures of cooperation than mail *response* rates, for which the denominator also includes units that turned out to be vacant or non-existent. This evaluation examines mail return rates at varying levels of geography, quantifies information about incoming questionnaires from occupied housing units according to form type and timing with respect to critical operational dates, and provides return rate data according to certain housing unit demographic characteristics.

B: Content and Data Quality

Overview

For Census 2000, the public will have five ways of providing census data. These modes include mailing back a questionnaire, filling out a census short form on the Internet, picking up and returning a Be Counted form, completing a census interview via telephone questionnaire assistance, or completing a personal visit interview with an enumerator. With this in mind, and the likelihood that the 2010 Census may offer additional options for response, studies in this category will document the characteristics of respondents and the mode by which they responded. Additionally, the data quality of each mode will be assessed. This category includes a Content Reinterview Survey study that will measure response variance, and a Master Trace Sample study. The latter will create a database containing a sample of census records with information pertaining to them from the entire census process. Other research will analyze the imputation process and evaluate multiple responses to the new race question.

What Will We Learn?

The findings from these evaluations will answer a number of critical questions on our process to define content (i.e., what questions to ask) and the resulting quality of data for Census 2000. These findings, in turn, can help us do a better job for Census 2010 and the American Community Survey.

We will learn about the completeness of the data by calculating item nonresponse rates and proxy response rates for all data items on the short and long forms. We also will look at demographic characteristics (such as age, sex, Hispanic origin, and race) of respondents by data collection mode and item nonresponse rates, and document the effects of data edit and imputation processes. We will assess responses to the new race question. In particular, we will recontact a sample of households with responses of two or more races, and ask each person to choose a single race category. This study is needed to meet the data requirements of other agencies that use only single race categories, and for comparison to 1990 Census race data.

We will also gain knowledge about data quality in comparison to external benchmarks by matching and comparing census data to data collected by the following Census Bureau surveys: Current Population Survey, Survey of Income and Program Participation, American Housing Survey, Residential Finance Survey, and the American Community Survey. The results of these matching and comparison studies will also help us to improve the design of future surveys and censuses.

(B.1) Analysis of the Imputation Process
To deal with missing data, three components will comprise the imputation process for Census 2000: substitution, edit, and allocation. Rates for each of these components will be produced for the 100 percent data items, for the tenure item, and for select sample housing unit and person items. This analysis will document the imputation process and will serve as a supplement to other evaluations.

(B.2) Documentation of Characteristics and Data Quality by Response Type
For Census 2000, there are five data collection modes available to respondents. Responses to the census may be collected from a mail or enumerator delivered census questionnaire; Simplified Enumerator Questionnaire used during nonresponse followup, list/enumerate, and update/enumerate; Be Counted form; Internet questionnaire; or reverse computer assisted telephone interview via Telephone Questionnaire Assistance. This evaluation will compare demographic differences and item nonresponse rates for these five response modes.

(B.3) Responses to Race Question
The purpose of this study is to create a data file for analytical purposes to allow comparisons between race data collected asking for only one race category and race data collected asking for two or more race responses. The study will include an oversample of households with responses of two or more race categories and ask each person to choose only one race category. This will allow us to measure the effects of this new question compared to the 1990 Census and will provide data needed by some government agencies that still require single race category data for historic comparability studies.

(B.4) Match Study of Accuracy and Coverage Evaluation Survey to Census 2000 to Compare Consistency of Race and Hispanic Origin Responses
The purpose of this evaluation is to determine the reliability of the race and Hispanic origin responses derived from the new race and Hispanic origin question. Reliability relates to the consistency with which responses for individuals are consistent across independent replications of the measurement process. This evaluation will be based on a comparison of census responses to those collected in the Accuracy and Coverage Evaluation survey.

(B.5) Content Reinterview Survey to Measure Accuracy of Data for Selected Population and Housing Characteristics
The Content Reinterview Survey utilizes a test-retest methodology, whereby a sample of households designated to receive the census long form are reinterviewed shortly after they have been enumerated by the census. These households are essentially asked the same question posed on the long form. Then the responses to the census and reinterview survey are compared. This survey assesses response variance and error that result from data collection and capture operations.

(B.6) Master Trace Sample
While most evaluation studies will provide detailed information on specific Census 2000 operations, the Master Trace Sample database will provide information that can be used to study the entire spectrum of operations, along with correlates of error across various systems, for a randomly selected group of census records. This database will contain, but is not limited to: address list information (e.g., source of address), final values for questionnaire items along with their values at each stage of processing, and enumerator information (e.g., number of enumerator attempts before completing an interview and enumerator production rates). This database also will contain information about the data capture system from rekeying and reconciling a subset of Master Trace Sample questionnaire images, the Accuracy and Coverage Evaluation, Content Reinterview Survey, and, possibly, administrative records.

(B.7) Match Study of Current Population Survey to Census 2000
Using the results of a person-level match of responses to the Current Population Survey (CPS) and Census 2000, this study provides a data set about differences between the Census and Survey estimates of social, demographic, or economic characteristics. Its strength is its ability to represent differences arising from non-sampling variation. The study focuses on the difference between CPS and Census estimates of poverty and labor force status (which are measured officially by the CPS) and on differences in reported race/ethnic status (which are measured quite differently on the two questionnaires).

(B.8) Comparisons of Income, Poverty, and Unemployment Estimates Between Census 2000 and the Current Population Survey
This study focuses on changes made to the Census 2000 questionnaire and forms processing systems that were designed to improve unemployment estimates. This evaluation examines whether these changes brought the Census 2000 unemployment estimates (for states, and for various demographic and socio-economic groups) closer to the official Current Population Survey estimates than they were in 1990. This analysis may be extended to compare data, definitions, and collection procedures with the Survey of Income and Program Participation.

(B.9) Housing Measures Compared to the American Housing Survey
In the past, the census and the American Housing Survey (AHS) have had a tendency to produce significant differences for many housing unit items. The purpose of this evaluation is to compare census housing data with data from the AHS and document the data differences between the two.

(B.10) Housing Measures Compared to the Residential Finance Survey
The census and the Residential Finance Survey (RFS) both collect similar housing information on mortgages, taxes, and insurance. However, the RFS collects this information in greater detail and directly from the files of mortgage lenders. This evaluation will compare housing data collected in the census with data from the RFS. Data to be compared include mortgage status, mortgage payments, presence of a second mortgage, second mortgage payments, real estate taxes, and property insurance payments.

(B.11) American Community Survey Evaluation of Follow-up, Edits, and Imputations
This evaluation will examine whether the content edit and followup procedures used in the American Community Survey have a measurable impact on any of the final estimates. This will be done by recalculating these estimates without the effects of the content edit and followup. These results will be used to infer the effect on decennial long form estimates of not conducting a similar content edit and followup procedure in Census 2000.

(B.12) Puerto Rico Race and Ethnicity
The methodology for this evaluation is being developed.

C: Data Products

Overview

The focus of this research is to determine the usability of selected data products and the effects of disclosure prevention measures on them. This will include studies of our data products strategy and of the Census Bureau's new electronic data dissemination system - the American FactFinder. We also will examine the limitations and effects of data swapping and our confidentiality edit – a combination of strategies used to prevent the disclosure of data that can be linked to an individual – on our data products.

What Will We Learn?

We will gain knowledge from these evaluations about the success of our data products strategy in meeting the needs of users and how we can improve it. We also will learn whether the American FactFinder is a usable and acceptable means to obtain census data. In studying our data swapping techniques, we will examine rates for different geographic levels and race groups and document new issues and problems that resulted from multiple responses to the race question.

(C.1) How Variations in Geography and Changes in Race Coding Affect Disclosure Prevention
For Census 2000, the data swapping methods first used in 1990 were refined through better targeting and expanded to include sample data. This evaluation examines variations in the effects of swapping due to: 1) a region's geographic structure, 2) a region's racial diversity, and 3) the number of dimensions used in the swapping.

(C.2) Usability Evaluation of User Interface With American FactFinder
The methodology for this evaluation is being developed.

(C.3) Data Products Strategy
The methodology for this evaluation is being developed.

D: Promotion and Partnership

Overview

During Census 2000, we will use new methods to promote census awareness and increase public cooperation. The primary goal of our comprehensive (and first ever) paid advertising campaign, coupled with an expanded partnership program, is to increase the mailback response rate, especially among historically undercounted populations. The advertising marketing strategy includes messages delivered through print media, radio, television, and out-of-home media (billboards, bus shelters, mobile billboards). The partnership program builds partnerships with state, local, and tribal governments, community-based organizations, and the private sector. Partners are asked to assist in three major areas: data collection support, recruitment, and promotion. In addition, a major school-based public information campaign will be launched to inform parents and guardians about the census through their school-age children. The planned evaluations for this research category will assess the effectiveness of these activities.

What Will We Learn?

These studies will help us understand how people's attitudes, knowledge, and behavior were affected by the paid advertising campaign. We also will compare these data to the 1990 census, which had no paid advertising campaign. We will examine which elements of the paid advertising media were reported/recalled most often by hard-to-enumerate groups, and provide data for Hispanics and for the five major race categories: African-American, Asian, American Indian and Alaska Native, Hawaiian/Pacific Islanders, and White. Specifically, we will look at what impact the marketing program had on the likelihood of returning a census form and (tentatively) whether it increased cooperation during nonresponse followup. The primary goals in studying the Partnership Program are to measure how well national and regional components accomplished their objectives in communicating a consistent census message of program initiatives and to determine which populations were best served by the program. Our assessment of the *Census in the Schools* program will tell us about the effectiveness of census educational materials and whether teachers receiving census materials incorporated them in their curricula.

(D.1) Promotion and Paid Advertising Campaign
The Census Bureau hired the National Opinion Research Center to conduct an assessment of the marketing and advertising campaign by fielding a survey before the campaign began and after the campaign has been launched. From this evaluation, we will assess intended and self-reported response behavior and will establish a baseline and pre- and post-census measures of awareness. We will obtain the actual response behavior for respondents to our survey. We will statistically model what effect self-reported advertising exposure has on the likelihood of responding to the census or cooperating with enumerators. This evaluation also will explore the link between raised awareness, knowledge, attitudes, and response to the census.

(D.2) Census in Schools Program
A post-census survey of school teachers will be conducted to assess the dissemination system for the Census in Schools materials and the effectiveness of the materials in motivating Census participation. Scholastic, Inc. will not conduct this evaluation as previously planned. We will hire an independent contractor to conduct the data collection activities for this program.

(D.3.) Mailout/Mailback Survey of Partners
This evaluation focuses on surveying participants in the Partnership Program by using a customer satisfaction questionnaire. We will assess the effectiveness of disseminating Census 2000 materials to partners, the types and value of in-kind services rendered, the specific partnership activities conducted, and the effectiveness of the program in reaching the hard-to-enumerate population. We also will obtain from non-Federal governments the financial demands placed on them as a result of Census 2000. The sample of partners will be selected using the Contact Profile and Usage Management System database. An independent contractor will be hired to conduct the data collection activities for this program.

E: Special Populations

Overview

The vast majority of U.S. residents live as families or individually in houses, apartments, mobile homes, or other places collectively known as “housing units.” However, there are millions of people in the United States who live in group situations such as college dormitories, nursing homes, convents, group homes, migrant worker dormitories, homeless shelters, or even in no place at all. Our evaluations will analyze the effectiveness of procedures to enumerate persons living in different types of group quarters and institutions. Some studies will focus on such things as enumeration at “service based locations” (shelters and food facilities for the homeless; outdoor locations where homeless people sleep). Major evaluations are planned for two operations designed to enhance the address list of special places: the Special Place Facility Questionnaire and the Special Place Local Update of Census Addresses.

What Will We Learn?

The findings from these evaluations will answer important questions on how effective enumeration procedures were in obtaining the count for group quarters. We will be able to analyze the completeness of the Facility Questionnaire and compare the telephone and personal visit operations. We also will assess interview completion rates for these groups and document the proportion of special population facilities that indicated if their administrative records could be made available. The evaluations will include distributions of these populations by type of group quarters, counts of persons at group quarters on Census Day who indicated a usual home elsewhere, and comparison of the predicted group quarters universe from the Facility Questionnaire operation with the group quarters universe as enumerated.

(E.1.a) Special Place/Group Quarters Facility Questionnaire - Operational Analysis
The Census Bureau's initial list of special places was supplemented by field staff and local partners during various operations (such as Block Canvassing and the Local Update of Census Addresses). This evaluation will document the number of special places added by each phase of this master list building operation. It also will document operational results and issues for the Computer Assisted Telephone Interview process used for most special places.

(E.1.b) Facility Questionnaire - Computer Assisted Telephone Interviewing and Person Visit
This evaluation will consist of personal visit reinterviews at a sample of special places to assess the accuracy of the information collected from the Facility Questionnaire via computer assisted telephone interview or personal visit. This evaluation will address whether classification discrepancies occur by type of special place and whether data quality differs by telephone or personal visit mode of data collection.

(E.2) Special Place Local Update of Census Addresses
This evaluation focuses on local governments' participation in the Special Place Local Update of Census Addresses. It will document changes to the address list along with operational issues that were encountered.

(E.3) Assess the Inventory Development Process for Service Based Enumerations
The purpose of this study is to assess the quality and effectiveness of the Service Based Enumeration (SBE) sites file. The quality of this file will be determined by the percentage of SBE records that were returned or could not be mailed because of incorrect addresses. We will look at deleted addresses, incomplete addresses, and added addresses. This study will also assess the addresses that could not be mailed by source to determine the relative merit of the various sources.

(E.4) Decennial Frame of Group Quarters and Sources
This study will evaluate the coverage, content, comparability, and the sources of information used to construct the group quarters frame for the decennial census (and American Community Survey), especially through comparison with the contemporary Business Register frame. This evaluation examines the feasibility and constraints to enrich or integrate these frames.

(E.5) Group Quarters Enumeration
This study will document various aspects of the group quarters enumeration. Some of the topics covered by this study include the total count of the group quarters population, the number of special places that were enumerated, and the number of group quarters that were enumerated. Additionally, the numerical distribution of group quarters per special place and of residents per group quarter will be documented.

(E.6) Service Based Enumeration

The goal of Service Based Enumeration (SBE) is to enumerate people without housing who may be missed in the traditional enumeration of housing units and group quarters. A complete enumeration of emergency shelters and soup kitchens, mobile food vans and non-sheltered outdoor locations will be conducted in late March 2000. This evaluation will document data collection completeness, last resort data collections, and whether the SBE unduplication process successfully identified individuals who were enumerated more than once. Also included in this study will be a profile of persons enumerated at SBE sites.

This page intentionally left blank.

F: Address List Development

Overview

These evaluations cover a broad spectrum of activities, both internal and external, involved with building address files and the related TIGER (geographic) database, including field operations from which address information and related map updates are gathered. The address list development category includes various evaluations of the Census Bureau's Master Address File (MAF), the TIGER database, and the Decennial Master Address File (DMAF). These include examination of the completeness and accuracy of address information in the MAF, as well as of the design of the MAF and DMAF. An evaluation of the U.S. Postal Service's Delivery Sequence File used in the MAF building process is also planned. A variety of census field and local/tribal partner operations will be evaluated to measure the impact of each operation on the MAF and the TIGER database. These include, but are not limited to: Address Listing, Block Canvassing, Update/Leave, List/Enumerate, multiple cycles of the Address List Review (also referred to as the Local Update of Census Addresses), and the New Construction Program. Combined, these field operations offer comprehensive address checks in rural and urban areas and are a primary source of address information used for MAF and TIGER database enhancement. Additional evaluations focus on the process of transferring address information to the MAF and incorporating map updates to the TIGER database from file sources and field operations.

What Will We Learn?

The findings from the address list development evaluations will provide insight into the most accurate methods for updating the MAF and the related TIGER database. This includes understanding the individual contribution of each operation as it is implemented. For each operation, we will look at the characteristics of addresses that were added, corrected, or flagged for deletion. We also will look at the geographic impact of each operation (i.e., we will examine how changes to the MAF are distributed geographically). Additionally, we will learn some things about the overall housing unit coverage in the census. Finally we will learn more about quality and coverage by examining addresses that are on the full MAF, but were not included in the census for various reasons. All of these evaluations will help inform continued MAF and TIGER database updating through the decade and also will provide insight for the 2010 Census and the American Community Survey.

(F.1) Impact of the Delivery Sequence File Deliveries on the Master Address File Through Census 2000 Operations

The Delivery Sequence File (DSF) is a file of addresses produced and maintained by the U.S. Postal Service. The Census Bureau uses this file, along with the 1990 census address list and other information, to create a permanent national address list called the Master Address File (MAF). For Census 2000, the Census Bureau will use the DSF as a primary source to enhance the initial MAF for mailout/mailback areas of the country. Subsequent DSFs will be used to update the address list through April of 2000, in order to maximize the inclusion of all existing addresses in the census. This evaluation will assess the impact of each of the DSFs through Census 2000 operations by profiling the number and characteristics of housing units added to and deleted from the MAF following each delivery of the DSF.

(F.2) Address Listing Operation and its Impact on the Master Address File

For Census 2000, an Address Listing Operation was used in update/leave areas of the country to create the initial Master Address File (MAF) and provide a comprehensive update of the streets/roads and their names in the TIGER database. In this operation, census enumerators went door-to-door to identify the mailing address and physical location of every housing unit as well as the existence and name of every street and road in areas where the U.S. Postal Service does not deliver mail using house number/street name addresses. The Census Bureau used this procedure in order to create a file of good locatable addresses for Census Bureau field operations in Census 2000 as well as its future demographic surveys, including the American Community Survey. This evaluation will assess the impact of the Census 2000 Address Listing Operation on the MAF by profiling the number and characteristics of housing units added to the MAF.

(F.3) Evaluation of Address List Review 1998

The Local Update of Census Addresses (LUCA) operation (also known as Address List Review) for Census 2000 included a LUCA 98 operation that focused on mailout/ mailback areas. For this operation, local and tribal government entities were provided a Census Bureau address list containing addresses derived from the Delivery Sequence File and the 1990 Address Control File. The objective of the LUCA operations was to provide local entities the opportunity to review the Bureau's address information and related maps and then provide feedback in the form of 1) address adds, deletes and corrections and 2) street and street name adds, deletions, and corrections on the maps. The Census Bureau compared the results to the block canvassing results in mailout/mailback areas, and all discrepancies were field verified. After Census Bureau review of submissions, local and tribal entities were given the opportunity to review results and to appeal situations in which they believed the Master Address File (MAF) still was incomplete or incorrect. This evaluation will assess the number and profile of housing unit adds to the MAF, the extent of geographic clustering of these adds, and the total number and profile of housing unit deletions and corrections. The evaluation also will include information documenting the participation rates of local and tribal governments and the proportion of addresses covered by these governments.

(F.4) Block Canvassing Operation and its Impact on the Master Address File
The objective of this evaluation is to determine the extent to which the Block Canvassing Operation corrected known Master Address File (MAF) deficiencies. In 1998, the Census Bureau conducted the MAF Quality Improvement Program which measured deficiencies in the MAF as it existed prior to the Block Canvassing Operation. These deficiencies included undercoverage, overcoverage, and geocoding errors. The Block Canvassing Operation is a dependent address updating operation conducted in mailout/mailback enumeration areas. This evaluation will assess the extent to which the Block Canvassing Operation has removed the deficiencies identified in the MAF Quality Improvement Program. That is, for the MAF Quality Improvement Program sample of housing units, we will examine the changes made during Block Canvassing to see if they are consistent with our expectations from that study.

(F.5) Block Canvassing Operation
For the 1990 census, the Census Bureau conducted an operation called Precanvass to improve its address list for mailout/mailback areas. For Census 2000, a similar operation, called Block Canvassing, was implemented. As with the1990 Precanvass, this operation was conducted primarily in areas where city-style addresses are used for mail delivery; however, for Census 2000, the Block Canvassing Operation covered a larger geographic area than did the 1990 Precanvass Operation, and the scope of the operation was expanded to include map (i.e. TIGER database) updates. The objective of this evaluation is to determine the overall effect of the Block Canvassing Operation on the Master Address File (MAF) by measuring the number and characteristics of housing unit adds, deletes, and corrections to the MAF.

(F.6) Address List Review 1999
The Local Update of Census Addresses (LUCA) operation (also known as Address List Review) for Census 2000 included a LUCA 99 operation for Update/Leave areas. For LUCA 99, local and tribal government entities were provided with census housing unit block counts that were created using addresses obtained from the Address Listing Operation. Participating entities were asked to review the counts and provide feedback when they believed the number of housing unit addresses for the block should have been higher or lower. Participating governments could challenge block counts, but could not provide specific housing unit adds, corrections, or deletes. Blocks that were challenged were sent to LUCA 99 Field Verification for relisting, then returned to participating governments for another review. This evaluation will document the participation rates of those tribal and local governments that were eligible to participate, the proportion of addresses covered by those governments, the number of blocks that were challenged and went to LUCA 99 Field Verification, and the extent to which changes occurred during the field verification.

(F.7) Criteria for the Initial Decennial Master Address File Delivery
In advance of the creation of the initial Decennial Master Address File (DMAF), address information was derived from a number of files and operations, particularly the 1990 Address Control File, the Delivery Sequence Files from the U.S. Postal Service, Block Canvassing, Address Listing and the Local Update of Census Addresses operations. The status codes from these files/operations were used

to determine which addresses from the Master Address File to include in the DMAF. This evaluation will provide a profile of these addresses as well as those MAF addresses not used for the DMAF.

(F.8) The Decennial Master Address File Update Rules

A number of address list update operations occur after the delivery of the initial Decennial Master Address File (DMAF) to the printing vendor. This evaluation will assess the profile of housing units corrected, flagged for deletion, and added to the DMAF from each update.

(F.9) New Construction Adds

In this new operation, the Census Bureau will request local and tribal governments to provide information on new construction addresses (i.e., addresses for units built after the Block Canvassing operation and not accounted for by subsequent Delivery Sequence File deliveries from the U. S. Postal Service), including street and street name updates to the maps. This evaluation will document the extent of local and tribal government participation and will document what happens to these cases during Census 2000 enumeration.

(F.10) Update/Leave

The Update/Leave operation is conducted in areas where mail delivery of questionnaires would be problematic. Field staff dependently canvass their assigned area, update the address list and map, and distribute a questionnaire to each housing unit. This evaluation will document address corrections, added units, and units flagged for deletion during the operation. We also will study problem referral forms completed by enumerators for difficult listing situations (e.g., unable to obtain access, gate blocked, road washed away, no trespassing signs), to see how well these situations were followed through on and how they might have contributed to coverage errors.

(F.11) Urban Update/Leave

Urban Update/Leave is an operation that targets whole census blocks and is conducted in areas where the Census Bureau is not confident that the addressed questionnaires will be delivered to the corresponding housing units. For Census 2000, 8 of the 12 Regional Census Centers have identified blocks for this operation. The Charlotte, Kansas City, Los Angeles, and New York Regional Census Centers decided to use tool kit methods exclusively in lieu of Urban Update/Leave. This evaluation will assess the number of addresses added and deleted as a result of Urban Update/Leave and will profile the housing unit addresses as follows: single/multi-unit, P.O. Box, drop/nondrop delivery, and LUCA 98 participant/nonparticipant. It also will examine the number of Urban Update/Leave addresses that match the Delivery Sequence File and the number and profile of addresses that result in a Master Address File correction.

(F.12) Update/Enumerate

Update/Enumerate is similar to Update/Leave, except that interviewers enumerate the unit at the time of their visit rather than leaving a questionnaire to be completed and mailed back. The operation is conducted in communities with special enumeration needs and where most housing units may not have house numbers and street name addresses. These areas include some selected American Indian

Reservations and the Colonias. Update/Enumerate also will be implemented in resort areas with high concentrations of seasonally vacant housing units. Most Update/Enumerate areas are drawn from address listed areas, but some may come from block canvasses areas. This evaluation will document the number and characteristics of housing units added, corrected, and flagged for deletion in Update/Enumerate areas.

(F.13) List/Enumerate

List/Enumerate is an all-in-one operation conducted in sparsely populated areas of the country. The address list is created and the housing units are enumerated concurrently. The main objectives of this evaluation will be to profile all addresses produced by the List/Enumerate operation, as well as to specifically profile the List/Enumerate addresses that matched to the Delivery Sequence File.

(F.14) Overall Master Address File Building Process for Housing Units

The objective of this evaluation is to examine the whole series of operations that affect the Master Address File (MAF) and the corresponding TIGER database during Census 2000 and to determine their individual impact on the final census inventory of housing units. This evaluation will assess 1) the effectiveness of each component of the Census 2000 MAF building process relative to the final list of housing units in Census 2000 and 2) which MAF update operations should be retained for MAF maintenance after Census 2000 is completed. It also will measure demographic and housing characteristics by operational source.

(F.15) Quality of the Geocodes Associated With Census Addresses

The objective of this evaluation is to measure the quality of residential address geocoding in Census 2000 and to identify the source of the geocode (i.e., the TIGER database, one of the several field operations, LUCA/New Construction participants, etc.).

(F.16) Evaluation of the Block Split Operation for Tabulation Purposes

Block Split operations are conducted by the Census Bureau to provide for tabulation of data where governmental unit and statistical area boundaries do not conform to collection block boundaries. This evaluation will measure the accuracy of block splitting operations for tabulation purposes.

This page intentionally left blank.

G: Field Recruiting and Management

Overview

Prompted by the difficulties in recruiting applicants and high turnover of employees in the 1990 decennial census, the Census Bureau redesigned its recruitment, staffing, and compensation programs for Census 2000. Several new programs were developed to address the 1990 issues and to help the Census Bureau successfully recruit several million applicants, hire several hundred thousand employees, and retain this staff through the decennial census. Some of these programs include frontloading, higher pay rates, and paid advertising.

What Will We Learn?

The purpose of these evaluations is to study the effects of these new program activities upon recruitment, staffing, and retention. A contractor, for example, determined that the 1990 District Office (now LCO) pay rates were not adequately set to attract and retain staff when compared to local economic conditions of that area. The methodology to set the Census 2000 pay rates, based on this knowledge, was revised and set to a derivative of the local prevailing pay rate. The effectiveness of this higher pay rate will be evaluated, as well as other recruitment and hiring programs (such as frontloading and paid advertising). In addition, this category contains an evaluation of the Operation Control System, a system used to track work going to and from field operations. We will also learn about the overall usefulness of this system.

(G.1) Census 2000 Staffing Programs
This evaluation examines the effectiveness of the Census 2000 hiring programs during Nonresponse Followup (NRFU). Study questions will focus upon the effectiveness of the higher pay rate program, frontloading, paid advertising, and other areas. Some of the questions are:
1) was the Census Bureau able to adequately hire and attract staff to execute NRFU, Accuracy and Coverage Evaluation, and other various field operations; 2) were the pay rates effective in attracting and retaining staff needed for Census 2000 NRFU; and 3) did recruiting activities provide an adequate supply of applicants and replacements. A portion of this study also will examine the effectiveness of the higher pay rates on productivity and evaluate the pay model as a predictor of local economic conditions.

(G.2) Operation Control System
This evaluation examines materials, such as debriefing questionnaires, management reports, systems documentation, and cost data, to assess the effectiveness of the Operation Control System (OCS) in tracking the cost and progress of field operations. This evaluation will answer the following questions: 1) was the OCS 2000 an effective tool for tracking work going to and from field operations; 2) were the products produced (for example, listings, labels, etc.) used in the manner for which they were designed; 3) was troubleshooting necessary during production and if so, was it effective; 4) did the management reports reflect production; 5) were the reports used in managing the operations; and 6) were Field Division's overall needs met with respect to the OCS 2000?

H: Field Operations

Overview

This category includes studies of various field operations and strategies whose goals are to curb questionnaire delivery problems and obtain census data from individuals who did not respond to the census by a specified date. For example, the Local Census Office (LCO) delivery of questionnaires returned by the U. S. Postal Service "as undeliverable as addressed" is designed to increase the number of questionnaires reaching potential respondents who may not have received one otherwise. The Nonresponse Followup operation consists of sending an enumerator to collect census data from every address from which no mail, telephone, or Internet response was received. Evaluations in this category will analyze whether these operations were conducted as planned and will assess their effectiveness. Additionally, operational results will be documented for each LCO for historical purposes.

Analyses in this category also will examine our efforts to count those categorized as hard-to-enumerate. 1990 Data for Census 2000 Planning, which was previously known as the Planning Database, is composed of 1990 person and housing unit census data that are indicators of nonresponse and potential to be undercounted. This database will help the Regional Census Centers determine the placement of Questionnaire Assistance Centers and Be Counted Forms. The database will also be used by participants of our partnership program. Studies in this category will evaluate the utility of the 1990 Data for Census 2000 Planning along with the usage of Questionnaire Assistance Centers. In addition, we will evaluate our targeted enumeration methods such as blitz enumeration (use of a group of enumerators to conduct enumeration in a compressed time frame), team enumeration (two enumerators working together where safety is a concern), and the use of local facilitators (long-time neighborhood residents or church leaders who assist the enumerator in gaining entry to the neighborhood).

Because some respondents will be able to provide data without a census identification number (e.g., Be Counted and Telephone Questionnaire Assistance), it is possible that respondents will submit addresses that are not on our Master Address File. We will conduct a field verification of these types of addresses. If an enumerator verifies that the address is a valid housing unit, then it will be added to the Decennial Master Address File. We also will conduct an evaluation of the effectiveness of this operation.

What Will We Learn?

The results of these evaluations will give us an indication of how successful we were at obtaining data from nonrespondents including the hard-to-enumerate, and how to better plan these types of operations for future censuses. The evaluation of Nonresponse Followup will report proxy rates, number of partial interviews, vacant rates, and number of units enumerated during final attempt procedures, which will help us to assess whether the operation was conducted as planned. Other analyses will provide information about the quality of our enumerator training program, the usefulness of the 1990 Data for Census 2000 Planning, and a profile of Local Census Offices which will contain various descriptive statistics.

Field Operation Evaluations

(H.1) Use of 1990 Data for Census 2000 Planning
For Census 2000, the Census Bureau has developed a Graphical User Interface that will work with the data in the 1990 Data for Census 2000 Planning database to aid Regional Census Centers in planning their specific operations. This evaluation will focus on the use of these data in census planning, which includes use by partnerships and identification of special advertising campaign areas, questionnaire assistance centers, and tool kit and Be Counted sites. In addition, the study will assess the geographic distribution of tracts targeted for said operations and sites.

(H.2) Operational Analysis of Field Verification Operation for Respondent Generated Questionnaires
Respondent generated questionnaires (e.g., Be Counted, Telephone Questionnaire Assistance) or questionnaires for which an enumerator is not able to verify that the address exists are referred to as non-ID housing units. During field verification, enumerators will visit the location of these non-ID housing units and verify their existence on the ground before they are added to the Decennial Master Address File (DMAF). For Census 2000, non-ID questionnaires that are geocoded to a census block, but do not match to an address already in the MAF will be assigned for field verification. This operational analysis will attempt to answer questions such as how many units were added to the DMAF after verification and if operational problems were encountered during the implementation of field verification.

(H.3) Local Census Office Delivery of Census 2000 Mailout Questionnaires Returned by U.S. Postal Service with Undeliverable as Addressed Designation
Due to a low mail response rate and a high Undeliverable as Addressed (UAA) rate during the 1990 Census, the Census Bureau will conduct a UAA delivery operation for Census 2000 and analyze how many UAA questionnaires were designated for delivery, and how many of these were successfully redistributed by the Local Census Offices (LCOs). This evaluation also will focus on those U.S. Postal Service Sectional Centers Facilities whose delivery area covered multiple LCO borders, and will determine if delivery was successful in those areas.

(H.4) Questionnaire Assistance Centers for Census 2000
The Census Bureau will provide walk-in assistance centers where respondents can receive assistance with completing their questionnaire. Language assistance guides will be available in over 40 different languages, along with Be Counted forms that will be available in English and five other languages. This study will document various aspects of the Questionnaire Assistance Centers (QACs) such as choice of location, hours of operation, and number of employees. In addition, the frequency of use of the QACs will be analyzed.

(H.5) Nonresponse Followup for Census 2000
This operation will be conducted for all housing units in the mailout/mailback and update/leave areas for

which the Census Bureau has not checked in a questionnaire by
April 11, 2000. During Nonresponse Followup (NRFU), enumerators will visit each nonresponding unit to determine the occupancy status of the unit on Census Day and to collect the appropriate data (i.e., long form or short form) for the household members. The objective of this analysis is to document various aspects of the NRFU operations. Some of the topics covered in this study include determination of NRFU workloads, identification of the demographics of those enumerated in NRFU, and documentation of the number of NRFU Simplified Enumerator Questionnaires that were partial interviews, refusals, completed via proxy respondents, or completed during final attempt procedures. The percent of NRFU units classified as occupied, vacant, or delete will be documented. Additionally, this evaluation will determine when each Local Census Office (LCO) started and completed their NRFU operation and the LCO cost of the operation.

(H.6) Operational Analysis of Non-Type of Enumeration Area Tool Kit Methods
Tool kit methods are special enumeration procedures (e.g., blitz enumeration, and the use of local facilitators) available for improving cooperation and enumeration in hard-to-enumerate areas. For this operation, the Census Bureau will assess the characteristics of areas targeted for tool kit methods based on the 1990 Data for Census 2000 Planning, and how often tool kit methods were used in areas not identified by these data.

(H.7) Evaluation of Nonresponse Followup Enumerator Training
During Census 2000, we will hire over 500,000 people to fill temporary positions. The largest number of these workers will be hired for the Nonresponse Followup (NRFU) operation. Adequate employee training will be critical to the success of NRFU. The overall objective of this evaluation is to examine the quality of the NRFU enumerator training program as well as the enumerator's state of preparedness following training. In addition, use of training materials and the adequacy of coverage of job assignments will be evaluated.

(H.8) Operational Analysis of Enumeration of Puerto Rico
Census 2000 is the first time that an update/leave mailback methodology will be used to conduct the enumeration in Puerto Rico. This evaluation will determine how many addresses were encompassed by this enumeration methodology, a profile of the addresses, and what operational problems were encountered in the field as a result of address list compilation and processing procedures. This study also will make comparisons to stateside Update/Leave data.

(H.9) Date of Reference for Respondents of Census 2000
The Census 2000 questionnaire states that the respondent should report age as of April 1, 2000. This study will document the average date of reference used by census respondents and the average date of reference by method of enumeration. This study also will document various types of discrepancies between date of birth and reported age. In addition, reported age and birth date on the census questionnaire will be compared to the same information collected by the Content Reinterview Survey.

(H.10) Local Census Office Profile

This operational summary will provide descriptive statistics at the Local Census Office (LCO) level for many census operations. For example, total housing units, average household size, and mail return rate will be among the statistics reported for each LCO.

I: Coverage Improvement

Overview

The coverage improvement evaluations examine various Census 2000 operations that are intended to improve the coverage of both housing units and people in the census. Following the mailback efforts to complete the census, a series of operations are conducted to ensure that people were counted at their correct Census Day address, confirm the status of housing units that were deleted or enumerated as vacant, and to ensure the inclusion of all persons in a household when the returned form shows discrepancies in the number of persons enumerated.

What Will We Learn?

From these evaluations we will learn about the effectiveness of these various operations as they attempt to improve census coverage. From the Nonresponse Followup operation, we will examine the potential coverage gain from identifying movers and checking to see if they were counted at their Census Day address. We will also analyze the situations where entire households were identified as having a "usual home elsewhere." For the Coverage Improvement Followup, we will examine the person and housing unit coverage gains from this operation, which determines the Census Day status of certain types of housing units (most of which are identified as deletes or coded as vacants in earlier census operations). The evaluation of the Coverage Edit Followup will measure coverage gains from this operation, which consists of contacting households whose completed forms show discrepancies regarding the number of persons enumerated, or whose completed form indicates there are more than six persons in that household. Furthermore, we will evaluate the coverage questions on the enumerator questionnaire to determine how well enumerators asked these questions and used the answers to obtain an accurate household roster.

Coverage Improvement Evaluations

(I.1) The Coverage Edit Followup for Census 2000
The Coverage Edit Followup (CEFU) is used to increase within household coverage and improve data quality in two ways. A standard questionnaire only has room for six persons, so CEFU is used to collect data on additional persons in large households. Second, it resolves discrepancies on mail return forms between the reported household size and the actual number of persons for which data are recorded on the census form. An attempt will be made to resolve all households that fail edits for these situations by using a Computer Assisted Telephone Interview. This analysis will document the workload, operational aspects, and coverage gains from conducting this operation.

(I.2) The Nonresponse Followup Whole Household Usual Home Elsewhere Probe
During the Nonresponse Followup (NRFU), List/Enumerate, and Update/ Enumerate operations, enumerators will ask respondents whether their address is a seasonal or vacation home and if the whole household has another place where they live most of the time. When respondents indicate they had a usual home elsewhere on Census Day, enumerators will record census information about this on a blank Simplified Enumerator Questionnaire (SEQ - a version of the mail return questionnaire that is easier to use for personal visit enumeration) and enumerate the current address as a vacant unit or obtain information about the people living there on Census Day. This evaluation examines how often SEQs were completed as Whole Household Usual Home Elsewhere (WHUHE), how many of these addresses were matched to an address on the Decennial Master Address File (DMAF) , how often addresses could neither be matched to the DMAF or geocoded, how often the WHUHE persons were already included on the census form for this address, and how often we found a different set of people on the census questionnaire for this address.

(I.3) Nonresponse Followup Mover Probe
In Census 2000, in-movers (households that moved there after Census Day) will be identified during the Nonresponse Followup (NRFU), List/Enumerate, and Update/Enumerate operations and will be asked if they were enumerated at their Census Day address. If a respondent does not recall completing a census form at their Census Day address, the enumerator will complete a questionnaire for the in-mover household using their Census Day address. This evaluation looks at how many of these cases occurred, and how many persons were added to the census as a result of this procedure.

(I.4) The Coverage Improvement Followup
The Coverage Improvement Followup (CIFU) universe will consist of units classified as vacant or deleted in NRFU, adds from the new construction operation, late adds from Update/Leave, blank mail returns, and lost mail returns. During CIFU, enumerators visit these units to verify the Census Day status and collect person and housing unit data as appropriate. This evaluation will document the person and housing unit coverage gain from conducting the CIFU, including the number of units that changed status from vacant to occupied or from delete to either vacant or occupied. This study also will examine the characteristics of persons and housing units added as a result of the CIFU, start/finish

dates, and the cost of the operation.

(I.5) Coverage Gain from Coverage Questions on Enumerator Completed Questionnaire
In 1990, enumerators began their interview with an explanation of who should be included as residents of the household. This procedure was changed for Census 2000. Now, enumerators begin by asking how many people were living or staying in the housing unit on Census Day. After collecting appropriate person and housing unit data, the enumerator will ask two coverage questions. The first asks about typical situations in which persons who should be included as residents tend to be missed – babies, foster children, persons away on business or vacation, roomers or housemates, and temporary residents with no other home. If someone has been missed, then he or she will be added to the form and their census information will be collected. The second question asks about typical situations in which persons who should *not* be included as residents tend to be included as such – persons away at college, in the armed forces, in a nursing home, or in a correctional facility. If someone was included on the form but should be counted elsewhere, then the enumerator will delete them from the form by marking the cancel box under their name. The purpose of this analysis is to study the effectiveness of the new coverage questions in the identification of persons who would have otherwise been missed or included in error.

This page intentionally left blank.

J: Ethnographic Studies

Overview

These evaluations will study certain aspects of coverage for various populations and attempt to identify areas where methods of collecting census data for these populations can be improved. We will examine within-household undercoverage, rostering methods (used to determine the residents of a unit), and household composition by comparing census results to Current Population Survey information for the same addresses. Another study in this category will apply social network field and analysis methods to evaluate census coverage and processes. We also will conduct ethnographic research on mobile populations and Colonias – areas lacking basic infrastructure and services along the border between the United States and Mexico.

What Will We Learn?

The comparison of the Current Population Survey and the census will give us greater insight into within household coverage errors by identifying who is missed in this survey by race, age, ethnicity, sex, and relationship. We will learn whether characteristics of the household (e.g., tenure, composition) are predictors of coverage. Other results will help us determine whether individuals can be better identified from their position in social networks (based on their interactions and transactions with others) than by comparing sets of address and person records. We will also learn how to improve procedures to enumerate mobile populations by tracing Census Day travel routes or stopover sites for a sample of such persons and determining undercounts or multiple enumerations of them in the census. We also will learn how to overcome barriers to enumerating Colonias in future censuses.

Note: Apart from the Census 2000 Evaluation Program; the Accuracy and Coverage Evaluation (A.C.E.) Survey for Census 2000 will provide a great deal of information pertaining to the coverage of various population groups. The A.C.E. Survey itself will be studied as part of the Census 2000 Evaluation Program in evaluation categories N, O, and P.

Ethnographic Studies

(J.1) Coverage, Rostering Methods and Household Composition: A Comparative Study of the Current Population Survey and Census 2000

Previous Census Bureau studies suggest a need for more analysis of the interactions between household composition, unusual living situations, and type of data collection (i.e., a census versus a survey). The census and CPS both collect household rosters and relationships information, and the CPS collects additional information on other relationships within households that can be used to study census data. This evaluation examines these within household coverage and household composition differences.

(J.2) Ethnographic Social Network Tracing

This study will use ethnographic and social network methods to study two questions: 1) What is the increased risk of coverage errors (both misses and erroneous inclusions) for persons who move once or more in a 6 month period around Census Day; and 2) Can these and similar coverage errors be reduced by identifying and tracing persons through their social networks and interactions. This would be a new approach for the Census Bureau, which traditionally has matched on non-household records (such as drivers license lists) to the census to identify persons at risk of coverage errors. The study also will examine how difficult it would be to incorporate social network methods into the census enumeration process. Various demographic, housing, household, social, and economic data will be documented for those persons who were missed or erroneously enumerated in the census.

(J.3) Comparative Ethnographic Research on Mobile Populations

In this study, a sample of selected mobile people will be traced to identify their Census Day travel routes or stopover sites. The information then will be matched and reconciled with census results. Coverage errors found in the census then will be analyzed to develop recommendations for improving procedures.

(J.4) Colonias on the U.S./Mexico Border: Barriers to Enumeration in Census 2000

Colonias are unincorporated, generally low income residential subdivisions lacking basic infrastructure and services (e.g., paved roads and public water systems) along the border between the U.S. and Mexico. In order to develop appropriate enumeration procedures and effective outreach and promotion programs for Colonias, it is necessary to better understand the unique situations and issues associated with conducting the census or other Census Bureau surveys in these areas. This research will examine the potential barriers to census enumeration in Colonias in the context of Census 2000 through participant observation, in-depth interviews, and focus groups with selected Colonia residents. Based on previous research, topics of particular interest include irregular housing, concerns regarding confidentiality, complex household structure, knowledge of English, and literacy.

K: Data Capture

Overview

The Data Capture System for Census 2000 (DCS 2000) will process more than 120 million census forms by creating a digital image of each page and interpreting the entries on each image using Optical Mark Recognition (OMR), Optical Character Recognition (OCR), or keying. These evaluations are designed to assess these components of DCS 2000, the Data Capture Audit Resolution (DCAR) process, and to measure the impact of each on data quality and on subsequent data coding operations.

What Will We Learn?

Findings from these evaluations will determine the level of accuracy at which the data capture system performed and how census data quality compares to that for capture systems used in 1990. Detailed information about the system will be collected, ranging from the number of forms processed by form type, date, and processing office, to measuring the accuracy of each of the three capture modes - OMR, OCR, and Key From Image. Operational problems and their resolution will be documented. Evaluation of the DCAR process will examine the system's ability to identify and resolve capture problems stemming from problems with response entries. Additionally, an evaluation of the interaction between the redesigned questionnaires and the new data capture system will be conducted. The impact of data capture errors on our ability to correctly assign industry and occupation codes will also be assessed.

(K.1.a) The Data Capture Audit Resolution Process
This evaluation documents the results of Data Capture Audit Resolution by failure reason, form type, and Data Capture Center. Using these same categories, it also will document the number and types of changes that can be made by Audit Review clerks and the results of the Audit Count review.

(K.1.b) Quality of the Data Capture System and the Impact of Questionnaire Capture and Processing on Data Quality
This evaluation examines how the data capture system affects data quality and whether the rules for determining where cases are routed (e.g., to key from image, Data Capture Audit Resolution or audit resolution) are set appropriately. In addition, this evaluation will document and compare the data quality of each data capture method for every field on the questionnaire, as well as by form type, Data Capture Center, and racial and ethnic categories.

(K.1.c) Analysis of Data Capture System 2000 Keying Operations
This evaluation will study various aspects of the Key From Image (KFI) and Key From Paper (KFP) operations. We will document the number of questionnaires processed and production keying rates in each of these operations by form type, Data Capture Center, cluster, and date. This study also will look at the accuracy of the data captured by the KFI operation in conjunction with OCR reject rates, the content distribution of fields accepted by OCR, and those rejected and sent to KFI. Our ability to recover from KFI/KFP operational problems, the adequacy and timeliness of management reports, the cost associated with the keying operation and our ability to hire and retain keying staff also will be assessed.

(K.1.d) Synthesis of Results from K.1.a, K.1.b, and K.1.c
This report will combine and summarize the results from the following studies: The Data Capture Audit Resolution Process (K.1.a), Evaluation of the Quality of the Data Capture System and the Impact of Questionnaire Capture and Processing on Data Quality (K.1.b), and Analysis Data Capture System 2000 Keying Operations (K.1.c).

(K.2) Analysis of the Interaction Between Aspects of Questionnaire Design, Printing, and Completeness With Data Capture
This study will focus on what impact the redesigned paper questionnaires used for Census 2000 had on respondent behaviors and on the ability of the new data capture process to completely and accurately convert the questionnaire data to computer files.

(K.3) Impact of Data Capture Errors on Autocoding, Clerical Coding and Autocoding Referrals in Industry and Occupation Coding
The information provided by respondents to the industry and occupation questions on the census form must be assigned (coded) to a standard set of categories. This evaluation examines how data capture errors affect the ability of the autocoding system and clerical coders to assign correct Industry and Occupation codes.

L: Processing Systems

Overview

Once census data from all sources are captured by the Data Capture System 2000, they are stored in a file known as the Decennial Response File (DRF). Several processes then must be applied before the data can be used to produce official census counts and tabulations. One process is applied to link multiple questionnaires that were used to enumerate that same household. For example, a large family could have a mail return form with data on six members of the household and an enumerator form with data on the rest of the household. Another process is used for situations where multiple questionnaires involving different households were received for the same address. For example, one form could be for a household that moved out near Census Day, and the other form could be for the household that then moved in. A computer program known as the Primary Selection Algorithm (PSA) then is used to decide which person and housing unit data should be used for census tabulations. Following all these processes, the DRF is merged with the Decennial Master Address File (DMAF) to create the Census Unedited File (CUF), which contains the original responses for a household.

A variety of post-census activities are needed to prepare the data from the original responses to releasing the official counts and tabulations. These activities include editing and imputation, coding of write-in response items (such as race, language, industry and occupation, and place of work/migration), tabulation recoding, and data disclosure avoidance.

The Beta Site is a software testing site for Bureau of the Census application developers and is used as an integration center for Regional Census Centers (RCC) and Local Census Offices (LCO) systems, a testing center for all systems, and a support center for RCC, LCO, and the National Processing Center systems. We will examine the effectiveness of this software testing site.

What Will We Learn?

Analysis of a reinterview of multiple questionnaire addresses will determine if the PSA methodology and rules for resolving these cases accurately identified the Census Day household members. The evaluation of the DRF creation and processes will examine how well multiple forms for the same household were linked. Analysis of CUF creation will document the number of times each specific DMAF/DRF rule was applied. The Beta Site analysis will include information on whether the data collection systems were successfully integrated, and the benefits of the software testing and release process.

(L.1.) Invalid Return Detection
The objective of this evaluation is to look for large geographic areas that may have high rates of multiple Be Counted forms and forms completed by telephone per housing unit.

(L.2) Decennial Response File Stage 2 Linking and Setting of Expected Household Population
This evaluation will document how frequently census forms were linked during the Decennial Response File processing and the types of linkages that were constructed. It will also assess the accuracy of the automated process for setting the expected household size and its effects on the census population.

(L.3.a) Analysis of Primary Selection Algorithm Results (Operational Assessment)
The objective of this evaluation is to document the effects of using the Primary Selection Algorithm in resolving situations when multiple household questionnaires are received for the same address.

(L.3.b) Resolution of Multiple Census Returns Using Reinterview
The objective of this evaluation is to determine the accuracy of Primary Selection Algorithm rules for determining the Census Day residents for an address. The data will be collected using a reinterview of a sample of respondents.

(L.4) Census Unedited File Creation
This evaluation documents the results of the process of determining the final housing unit inventory. The final housing unit inventory for the census is determined during the process of creating the Census Unedited Detail File. The final housing unit inventory is created by merging information on the processed Decennial Response File with the information on the Decennial Master Address File.

(L.5) Beta Site
This evaluation will answer questions about how well the Beta Site integrated the data collection systems, and its overall utility for software testing and release.

M: Quality Assurance Evaluations

Overview

For Census 2000, the overall objective of the Quality Assurance (QA) program is to assist in producing deliverables or outputs which meet the Bureau's quality requirements. The QA program will identify when major inputs such as people, material, machinery, software, etc. do not meet quality requirements. The QA data will provide managers, supervisors or employees with information to make necessary adjustments and improvements to the system. At the end of the operational task, QA will identify and correct clusters of outputs which contain a significant number of errors.

What Will We Learn?

The QA evaluations will provide information to help determine if the QA approach used in Census 2000 is the right approach in a census environment, whether the QA operation improved the overall quality of the census, how effectively it was implemented, and how it might be improved. For example, the results of the first study will help us determine if different QA approaches should be explored for census use. For the second study, the effectiveness of variables that are used to detect enumerator falsification will be measured, and appropriate variables will be added and/or deleted from the detection process.

(M.1) Assessment of the Quality Assurance Approach During Census 2000
For Census 2000, the overall objective of the Quality Assurance (QA) program is to assist in producing deliverables or outputs that meet the Census Bureau's quality requirements. To achieve this objective, the QA program, whenever possible, focuses on three main concepts or philosophies: prevention, improvement, and protection. The goals of this study are to document operational experiences with this approach in Census 2000; measure quality levels that were achieved, and determine if other approaches should be explored for Census 2010.

(M.2) Effectiveness of Existing Variables in the Model Used to Detect Fabrication During Reinterview, and the Identification of New Variables
The reinterview program is a quality assurance operation whose major objective is to detect enumerators who may have falsified data. This evaluation examines variables used in the fabrication model to determine if they were effective in detecting fraud; whether other variables should be added to the model; and to provide suggestions on other ways to improve this program.

N: Accuracy and Coverage Evaluation Survey Operations

Overview

The Census Bureau will conduct the Accuracy and Coverage Evaluation (A.C.E.), a nationwide sample survey, to determine the number of people and housing units missed or incorrectly counted in the census. The basic approach is to independently relist a sample of blocks, re-enumerate them during the A.C.E. survey, and then compare the results to the census data for the same blocks. The Census Bureau will use the results of the A.C.E. to correct the census counts obtained through the preceding enumeration procedures.

The studies in this category will measure how well the Census Bureau carried out different components of the A.C.E. For instance, analysis projects and evaluations will be conducted that measure the completeness of the housing unit lists used for A.C.E. interviewing, the quality of the A.C.E. person interviewing process, and the accuracy of the procedures used to match persons counted during the A.C.E. interview to those that were enumerated in the census. The success of each A.C.E. component affects the quality of the final estimates.

What Will We Learn?

The results of these A.C.E. analysis projects and evaluations will help the Census Bureau to document this coverage measurement operation and improve its procedures. For example, we will learn whether match rates were different in relisted blocks. An examination of laptop computers used during person interviews will identify errors encountered by interviewers and also will provided suggestions for how to improve the computer assisted instrument in the future. Other studies will determine how well we detect interviewer fabrication, while also looking at its effect on A.C.E.

These operational analyses and evaluations will document the A.C.E. process and give the Census Bureau greater insight into what causes error in the measurement of coverage error. Some causes of error are attributable to census questionnaire data capture. Moreover, matching errors may add to errors in the estimates of census coverage. One evaluation in this category will examine a subsample of rematched A.C.E. blocks to measure matching errors. We also will measure the effect of matching error on Dual System Estimates and undercount rates.

The evaluations in this category will help the Census Bureau to identify operational causes of error in measuring coverage and will help to minimize them when planning future censuses.

(N.1) Contamination of Census Data Collected in A.C.E. Blocks
This evaluation examines whether census and A.C.E. operations were kept operationally independent (a key requirement for avoiding bias in the dual-system estimates of coverage error) by comparing census results in A.C.E. and non-A.C.E. clusters.

(N.2) Analysis of Listing Future Construction and Multi-Units in Special Places
A new procedure during block relisting for A.C.E. 2000 is to include housing units under construction, and multi-unit structures within special places. We will study the effectiveness of the listing, the number of units added to A.C.E., and match rates to the census.

(N.3) Analysis of Relisted Blocks
The A.C.E. address listing operation is reviewed for high concentrations of geocoding errors, and blocks with too many errors are relisted. This analysis will examine the relisted blocks to document their characteristics and the results of matching to census listings.

(N.4) Analysis of Blocks With No Housing Unit Matching
For most blocks in the A.C.E., the matching is done in two steps: first the addresses are matched to the census address list, and then the persons are compared to the census questionnaires for matching addresses. The housing unit match step is not done for relisted blocks (see N.3) or for blocks in list/enumerate areas. The purpose of this study is to determine how this affects the person matching process.

(N.5) Analysis of Blocks Sent Directly for Housing Units Followup
The A.C.E. addresses first are computer matched to the addresses on the January 2000 version of the Decennial Master Address File and then undergo a clerical matching operation. Some blocks are sent directly for a follow-up interview when there is little perceived benefit to clerical matching. While this allows the field follow-up to begin earlier, it also may reduce the ability of that operation to resolve the match status of these units. This study will examine the effectiveness of this strategy by comparing match rates and unresolved rates for these blocks to those for blocks that did undergo clerical matching.

(N.6) Analysis of Person Interview With Unresolved Housing Unit Status
This analysis examines whether housing units with an unresolved status after the initial housing unit match are eventually resolved during the person interview and final housing unit match operations.

(N.7) Analysis on the Effects of Census Questionnaire Data Capture in A.C.E.
During the A.C.E. person matching, data capture *images* of census questionnaires are examined when the initial match results indicate the census data are insufficient for matching. This study will document how often this occurred and the effects on final match codes.

(N.8) Analysis of the Census Residence Questions Used

During A.C.E. interviewing and field followup for non-matches, persons are asked about their Census Day address and the results are used to determine where they should have been counted in the census. Persons counted at the wrong address then are classified as erroneous enumerations by the census. This study will examine the responses to these residency questions during A.C.E. and document how they affected the estimates of erroneous enumerations.

(N.9) Analysis of the Person Interview Process

This study examines the overall interviewing process. The analysis will include topics such as Computer Assisted Personal Interview instrument (i.e., laptop) performance, detection of interviewer falsification by the interviewing quality assurance process, and the impact of allowing interviewers to modify address information during matching operations.

(N.10) Falsification in A.C.E.

This evaluation examines how well the quality assurance process identified interviewers who entered false data in the A.C.E. interview and the impact of undetected false data on A.C.E. estimates.

(N.11) Extended Roster Analysis

During the census, an extended roster is used to capture names (but not data)for people in large households. A follow-up operation then collects demographic data for these people. If the follow-up does not collect the data, it will not be possible to match A.C.E. data for these people to census results and sending this case to A.C.E. field followup is pointless. This study will document how reviewing these extended rosters affected the A.C.E. person matching and followup operations.

(N.12) Matching Stages Analysis

The person matching is conducted first by computer and then undergoes three levels of clerical matching by clerks, technicians, and analysts. The goal of this analysis is to document the differences in the match codes assigned by these four different operations.

(N.13) Analysis of Unresolved Codes in Person Matching

Results from the Census 2000 Dress Rehearsal Person Followup interview indicated that there were a large volume of cases coded as unresolved when the interview was conducted with a proxy respondent. In general, proxies were able to answer whether a household lived at a given address on Census Day, but answered “Don’t Know” to questions regarding a household being in a group quarter and/or having a ususal home elsewhere, which resulted in an unresolved code. The goal of this analysis is to document the coding results for specified patterns of “Don’t Know” answers from proxy respondents. The coding results of proxy respondents will then be compared to those with similar patterns of answers from actual (i.e., non-proxy) respondents.

(N.14) Evaluation of Matching Error
A potential source of error in the coverage estimates are the matching operations used to classify persons as missed or erroneously enumerated in the census. This evaluation will determine the relative error associated with the matching operations and how matching error affects the Dual System Estimates.

(N.15) Outlier Analysis in the 2000 A.C.E.
In 1990, an outlier review was conducted in 104 of the blocks that contributed most to the net undercount. This review was conducted in 1991 after all operations were completed. In 2000, the outlier review is planned to be conducted before the matching is completed. Blocks will be selected for an indepth review by the analysts. Matching errors will be corrected, if they exist. In addition, the analysts will document the results of their investigation. This project will document the outlier review.

(N.16) Impact of Targeted Extended Search
This evaluation has two main purposes. The first study looks at the nature and extent of errors resulting from limiting the search area to one ring of blocks (around the sample block clusters). This is accomplished by looking at the effect of this Targeted Extended Search on the Dual System Estimates and variances for the evaluation post strata, as well as data from the production matching operation. The second study evaluates potential gains from adding a second ring of blocks to the search and match operation.

(N.17) Targeted Extended Search Block Cluster Analysis
In 1990, the search area for matching was extended to surrounding blocks for all clusters. In 2000, this only will be done for clusters deemed most likely to benefit this additional searching. This study will document the characteristics of such blocks and the effects of this strategy on final match rates.

(N.18) Effect of Late Census Data on Final Estimates
The aim of this evaluation is to determine the effect on A.C.E. adjustment factors of ignoring the small amount of census data collected after late September.

(N.19) Housing Unit and Person Coverage Analysis
This analysis provides an overall assessment of the quality of housing unit and person coverage A.C.E. operations. Some of the topics addressed in the analysis are quality of A.C.E. listing, effect of housing unit followup interviewing on the enhanced list, effectiveness of housing unit and person followup quality assurance, and noninterview rates.

(N.20) Group Quarters Analysis
In 1990, sample interviews were conducted in noninstitutional and nonmilitary group quarters. In 2000, A.C.E. sample interviews will not be conducted in any group quarters. The A.C.E. sample interviews will only be conducted in housing units, but sometimes it is difficult to determine if a place is a housing unit or a group quarters. The A.C.E. sample nonmatches in whole households will be compared to the group quarters enumerations in the census. The purpose of this analysis is to document these matching

results, identify post-strata containing a large number of group quarters, and to examine whether definitional problems led to group quarters classification errors.

(N.21) Analysis of Mobile Homes

Mobile homes were missed at a higher rate in 1990 than single family homes. Mobile homes are not a housing category for Census 2000, but they will be identified in the A.C.E. sample. This study will document the nonmatch rates for mobile homes and people in mobile homes from the A.C.E. sample matching results.

This page intentionally left blank.

O: Coverage Evaluations of the Census and of the Accuracy and Coverage Evaluation Survey

Overview

The studies in this category include a group evaluating A.C.E. coverage and a group evaluating census coverage. These studies will identify person and housing unit characteristics that are related to being missed or erroneously enumerated. Analysis in this area will also study the quality of data from proxy respondents, and the frequency and patterns of geocoding error. Furthermore, census counts and dual system estimates will be compared to demographic benchmarks to evaluate accuracy and completeness.

What Will We Learn?

Results from these evaluations will allow us to determine how complete our Master Address File was for Census 2000. Net coverage rates of housing units will be computed at the national and subnational levels along with gross omission and erroneous enumeration rates. Other studies will explain factors that contribute to housing unit coverage error. For example, we will learn whether type of address (city style versus noncity style) has an effect on housing unit coverage. In addition, there will be a study of housing unit duplication, to identify characteristics of duplicate units and their operational source.

Similarly, we will identify factors that contribute to person coverage error. For instance, studies will examine how nonmatch rates are affected by type of enumeration area (e.g., mailout/mailback, update/leave) and characteristics of blocks, households, and people. We will acquire knowledge about erroneous enumerations by determining which demographic, housing unit type, and type of enumeration variables are associated with them. Furthermore, we will conduct an analysis of measurement error, which will help us determine why people are erroneously listed in the census and the Accuracy and Coverage Evaluation.

(O.1) Type of Enumeration Area Summary
The census is conducted differently in the different types of enumeration areas. This project will document nonmatch rates and erroneous enumeration rates in the different areas. Geocoding error in the different enumeration areas also will be documented.

(O.2) Coverage of Housing Units in the Early Decennial Master Address File
The initial housing unit matching is done between the A.C.E. housing units and the housing units on the January 2000 version of the Decennial Master Address File (DMAF). This matching is conducted to link the ACE and census housing units for later processing. If an A.C.E. housing unit is linked to a census housing unit, the telephone number from a mail return questionnaire can be used to start person interviewing by phone. The results of the matching will allow an early look at the quality of the DMAF in January 2000.

(O.3) Housing Unit Coverage on the Master Address File
This evaluation assesses 1) the net coverage rate of housing units, 2) the gross omission rate of housing units, and 3) the erroneous enumeration rate of housing units. These assessments are made at the national level, smaller geographic levels, and for each post-strata. This evaluation also examines the potential impact on housing unit coverage had we excluded specific Master Address File building operations. This study is similar to the Housing Unit Coverage Study conducted in 1990.

(O.4) Analysis of Conflicting Households
During A.C.E. housing unit matching, situations are found where the census and A.C.E. listed two entirely different families. This study will document the follow-up interviewing results for these households to determine if the census was in error, the A.C.E. was in error, if the two families both live at the address, if there was misdelivery of the census form, and so on.

(O.5) Analysis of Proxy Data in the A.C.E. and in the Census
Both the census and A.C.E. sometimes must collect data from proxy respondents--persons who are not members of the household where data are needed. This study will examine match rates and erroneous enumeration rates for such cases in both the census and the A.C.E.

(O.6) P-Sample Nonmatches Analysis
This study will examine nonmatch rates for the post-strata used to form final dual-system estimates of census coverage errors. It also will examine these rates for other variables not used to form post-strata.

(O.7) Analysis of Person Coverage in Puerto Rico
The measurement of person coverage, and evaluation studies of that measurement, will be done separately for Puerto Rico. This study will document those findings and compare the results to those from the A.C.E.

(O.8) Analysis of Housing Unit Coverage in Puerto Rico
The measurement of housing coverage, and evaluation studies of that measurement, will be done separately for Puerto Rico. This study will document those findings and compare the results to those from the A.C.E.

(O.9) Geocoding Error Analysis
A housing unit and its occupants are classified as geocoding errors by the census if that housing unit is enumerated within the ACE search area and should not have been. This study will examine the frequency of geocoding error and will identify operations more prone to making such errors.

(O.10) Housing Unit Duplication in the 2000 Census
Duplication in the census is one type of erroneous enumeration. This analysis will identify duplicate housing units in Census 2000 and their characteristics. The study will also determine if duplication is more likely for one group or another (e.g. owners vs. renters). The census operations most likely to produce housing unit duplication will be identified, along with the most plausible sources of duplication.

(O.11) E-Sample Erroneous Enumeration Analysis
This study will examine erroneous enumeration rates for the post-strata used to form final dual-system estimates of census coverage errors. It also will examine these rates for other variables not used to form post-strata.

(O.12) Analysis of Nonmatches and Erroneous Enumerations Using Logistic Regression
This project looks at logistic regression as a tool to analyze the A.C.E. data. This purpose of this analysis is to build logistic regression models that relate demographic, housing unit type, and type-of-enumeration variables, to census nonmatches, A.C.E. nonmatches, and erroneous enumerations.

(O.13) Analysis of Person and Housing Unit Data Combined
For some housing units, the A.C.E. results will identify both missed and erroneously enumerated persons. This study will examine the person and housing characteristics of such cases.

(O.14) Analysis of Measurement Error
Measurement error is the term used for error in surveys due to an inability to collect the correct answer to a question. Measurement error can be attributable to the interviewer, the respondent, or the data collection instrument. The main question of this evaluation is whether or not the errors in A.C.E. residency status codes and person match codes had a significant effect on the Dual System Estimates. Other topics examined by this evaluation are the magnitude of the error attributable to the respondents or interviewers in A.C.E. Person Interview (and the A.C.E. Person Followup Interview), and the characteristics of people whose person match and residency codes were found to be incorrect in A.C.E.

(O.15) Impact of Housing Unit Coverage on Person Coverage Analysis

This analysis will include an examination of the effect that census housing unit updating operations (e.g., postal check, Local Update of Census Addresses) have on person coverage. The study also will identify characteristics of persons who were missed or erroneously enumerated due to housing unit errors, and compare them to the characteristics of those who were missed or erroneously enumerated for other reasons.

(O.16) Person Duplication in the 2000 Census

People are duplicated in the census for many different reasons. This analysis will identify the number and characteristic of duplicate persons in the 2000 Census. The study will also determine if duplication is more likely for one group or another (e.g., owners/renters). The census operations most likely to cause duplication will be identified, along with the most plausible sources of the duplication.

(O.17) Analysis of the 0-17 Age/Sex Post-Strata

Children have historically been disproportionately undercounted in the census. This study will examine coverage errors for children. It also will examine how often we encounter A.C.E. households in which all of the residents are under 16 years of age. These cases are treated as non-interviews and thus can affect the final coverage estimates for persons age 0-17. The characteristics of the people in these types of households and the household composition will be documented.

(O.18) Synthesis of What We Know About Missed Census People

The purpose of this study is to summarize and synthesize findings from A.C.E. and other sources about the causes and characteristics of persons missed or erroneously enumerated in the census.

(O.19) Analysis of Coverage of Housing Units in the Early Decennial Master Address File and Subsequent Census Coverage Improvement

The goal of this study is to assess the completeness of housing unit coverage on the early Decennial Master Address File (DMAF). We will determine which census operations contributed to undercoverage by deleting units that should have not been deleted, and which operations improved coverage by adding units not previously accounted for. We also will identify which census operations reduced housing unit duplication

(O.20) Consistency of Census Estimates with Demographic Benchmarks

This study uses independent demographic benchmarks to evaluate the accuracy of the Census 2000 counts and the completeness of coverage in Census 2000. While this approach cannot produce estimates for as many demographic groups and geographic areas as A.C.E., results can be compared to A.C.E. at aggregate levels.

(O.21) Implications of Net Census Undercount on Demographic Measures and Program Uses
This evaluation will address the effect of net census undercount on demographic measures such as growth rates, race, age and sex composition, and vital statistics rates used by a variety of programs and data users.

This page intentionally left blank.

P: Accuracy and Coverage Evaluation Survey Statistical Design and Estimation

Overview

The evaluations in this category examine the quality of Accuracy and Coverage Evaluation (A.C.E.) estimates. Analyses in this area will address the quality of Dual System Estimates (DSE) by examining estimates of variances and coefficients of variation. We will also analyze missing data procedures, compare A.C.E. results to various quality measures, and conduct a study of total error in A.C.E.

What Will We learn?

We will gain knowledge about the quality of A.C.E. estimates from the total error analysis, which will examine model and measurement error in the empirical DSE. We will also learn about the overall quality of A.C.E. by comparing its results to a synthesis of quality measurements from various coverage measurement evaluations and operational analyses. We will examine characteristics associated with missing data and the bias and uncertainty associated with the missing data procedures.

(P.1) Measurement of Bias and Uncertainty Associated With Application of the Missing Data Procedures
The purpose of this evaluation is to obtain a measure of the error (and sensitivity to assumptions) of missing data procedures used for the dual system estimates resulting from the A.C.E.

(P.2) Synthetic Design Research/Correlation Bias
Synthetic estimation uses a statistical model to modify coverage estimates for a particular post-strata using information from sample units outside the geographic area of interest. Because this can introduce bias, the accuracy of the Dual System Estimate (DSE) depends on the validity of the model and whether the assumptions of the synthetic model are satisfied. The purpose of this study is to determine if the assumptions were satisfied for Census 2000 A.C.E. post-strata, and to measure the effects of any biases on the DSEs.

(P.3) Variance of Dual System Estimates and Adjustment Factors
This study is designed to evaluate the quality of the Dual System Estimates (DSE) and adjustment factors by examining estimates of variances and coefficients of variation (CVs). We will compare the 2000 DSE variance estimates and CVs to the 1990 DSE variance estimates and CVs at the national-level and for various demographic variables.

(P.4) Overall Measures of A.C.E. Quality
The aim of this study is to synthesize quality measures from various coverage measurement evaluations and operational analyses to assess the overall quality of the 2000 A.C.E. Measures such as response rates, imputation rates, match rates, and correct enumeration rates by various demographic and geographic groups will be examined and, where possible, compared to the 1990 census rates.

(P.5) Total Error Analysis
The total error analysis will examine model and measurement error in the empirical Dual System Estimates. For each evaluation post-stratum, we will estimate the bias and variance in the net undercount rate for each type of nonsampling error and estimate the overall bias and variance in the net undercount rate.

Q: Organization, Budget, and Management Information System

Overview

Research in this category will document headquarters decision making processes and the impact of headquarters organizational structure on the decennial census. We plan to study the effectiveness of the Management Information System (MIS) and the Cost Model. The MIS is a data warehouse of cost and progress information for census operations that also includes an activity schedule for the decennial census. The cost model is used to formulate budgets for operations, allocate funds, and to assist in planning future census operations. Furthermore, we will conduct an evaluation that examines the role of contracting in carrying out Census 2000.

What Will We Learn?

The findings from these studies will help the Census Bureau to better manage future censuses and similar projects. These studies will document how well the MIS and the Cost Model worked in helping us manage Census 2000. We will compare the activities and recommendations of the 2000 research and development program to what was actually implemented for Census 2000 to determine which projects were most beneficial. In addition, we will examine the roles and influences of both external and internal entities on planning and implementing the census. Some of the groups that will be studied include various advisory committees, Office of the Inspector General, Congress, General Accounting Office, Census Bureau Monitoring Board, Census Bureau Executive Staff, and the Department of Commerce. Other research in this category will give us insight into the effectiveness of hiring contractors to help conduct Census 2000. We will learn how cost effective our contracts were and whether contractors are bringing in the expertise needed by the Census Bureau. We will also address whether the Census Bureau is losing "corporate knowledge" by giving contractors a major role in conducting Census 2000.

(Q.1) Management Processes and Systems of the 2000 Decennial Census
The purpose of this study is to determine how well various processes and systems worked for managing Census 2000. This analysis will include an evaluation of the Management Information System and the Cost Model. The effectiveness of decision making groups/processes (e.g., Census Operational Managers and decision memos) will be assessed. This study will also look at the organizational structure, roles, and influences of entities such as the Census Bureau Executive Staff, Department of Commerce, Inspector General, advisory committees, General Accounting Office, Census Bureau Monitoring Board, and Congress. Furthermore, a comparative study of management models will be conducted.

(Q.2) Effectiveness of the Contracting in Carrying Out the 2000 Decennial Census
The goal of this evaluation is to measure the effectiveness of contracting in Census 2000. We will look at the cost effectiveness of our contracts along with whether contractors are bringing in the expertise that is needed by the Census Bureau. This study will examine how well we managed our contracts and will determine if we are losing "corporate knowledge" by giving contractors a great deal of responsibility in this decennial census.

R: Automation of Census Processes

Overview

These studies will examine many of the major automated systems designed to support Census 2000. In general, we will assess whether the right requirements were defined for each of the systems and use this information to guide improvements needed for future censuses and surveys. The systems to be studied include:

- Telephone Questionnaire Assistance
- Internet Questionnaire Assistance
- Internet Data Collection System
- Operational Control System 2000 (OCS2000) System
- Preappointment Management System/Automated Decennial Administrative Management System (PAMS/ADAMS)
- American Fact Finder
- Data Capture
- Matching, Review, and Coding System
- Accuracy and Coverage Evaluation Survey 2000 System

What Will We Learn?

In addition to examining systems requirements, we will also assess other factors such as reliability and functionality, maintenance and security needs, and respondent acceptance. A common protocol will be designed to include general questions for the selected automated systems and debriefings will take place. We will also identify questions and concerns unique to specific systems.

Note: The plans for this evaluations category are currently being finalized. The specific evaluations will be identified in an updated version of this document.

March 29, 2000 DRAFT - Revised

Census 2000 Operational Analyses for Sampling and Estimation

A Brief Overview

Topics covered:

1. A.C.E. Sample Design
2. Long Form Sampling
3. Service Based Enumeration (SBE) Estimation
4. P- and E-Sample Missing Data
5. Dual System Estimation (DSE)
6. A.C.E. Variance Estimation
7. Long Form Estimation/Variance Estimation
8. A.C.E. Weight Trimming
9. Generalized Variances
10. Block Level Estimation
11. Unclassified Estimation
12. Housing Unit (HU) Dual System Estimation (DSE)
13. A.C.E. Targeted Extended Search

1. A.C.E. Sample Design

The A.C.E. is a multi-stage sample design. Initially, the original ICM sample of 750,000 housing units will be selected and be sent to the field to be independently listed. This listing sample is a state-based design and will include an oversample of large block clusters, a separate sample of small block clusters in each state. In addition, a separate sample of American Indian Reservation clusters will be selected. After the sample is listed, the A.C.E. sample cluster reduction and the small block subsampling will be done. After the initial housing unit matching operation and followup, the large block subsampling will be done to reach the final A.C.E. interview sample size. The E-sample will generally overlap with the P-sample.

Questions to answer:

- What are the sample sizes in terms of clusters and expected housing units?
- What are the unbiased sampling weights after each stage of selection?
- What are the weighted population distributions by State?
- What are the sample sizes (clusters and housing units) by TEA and ACERO?
- What's the average number of interviewed A.C.E. housing units in a cluster?
- What's the sample size (clusters and housing units) by stratum before and after sampling?

Processes included:

Initial Sampling
A.C.E. Reduction
Small block subsampling
Large block subsampling
E-sample Identification

2. Long Form Sampling

Long form sampling uses four sampling rates based on the size (MOS) of the long form sampling entity (LFSE). MOS is an estimate of occupied housing units. Long form sampling entities are geographic and statistical areas eligible for the sampling. Sampling rates are assigned to the collection blocks as follows:

- All blocks in Puerto Rico 1-in-6
- 1-in-2 for governmental units if estimated MOS is less than 800.
- 1-in-4 for government units if MOS is 800 or more but less than 1200.
- 1-in-8 for census tracts with MOS $\geq$ 2000.
- 1-in-6 for all remaining blocks in tract with MOS less than 2000.

Questions to answer:

- What are the sizes of the universe by sampling rates?
- What are sample sizes by type of enumeration area?
- What are the response rates for long form at different geographic level?
- What are the response rates for long form by selected household type?

3. Service Based Enumeration (SBE) Estimation

SBE is a fundamentally different approach to counting persons without a usual home than was used in the 1990 Census. SBE counts people at facilities such as shelters, soup kitchens, mobile food vans and certain outdoor locations. SBE estimation uses the multiplicity estimator methodology for the estimation of people with no usual residence who use SBE facilities. The multiplicity estimator depends on the service usage question for shelters, soup kitchens, and mobile food vans. SBE multiplicity estimation was done in the Census 2000 Dress Rehearsal. It will be used for Census 2000 for the redistricting file but not for the apportionment counts.

We will use multiplicity estimation to estimate the number of persons that use SBE facilities and do not have a usual home.

Questions to answer:

- How many persons were enumerated as part of SBE on Be Counted Forms (BCFs) and at shelters, soup kitchens and mobile food vans, and Targeted Non-Shelter Outdoor Locations (TNSOLs) after unduplication (i.e., apportionment count)?

- What is the distribution of response and nonresponse to the usage questions by nonresponse adjustment strata? What is the average respondent usage? (Separate for shelters and soup kitchens).

- For soup kitchens, what is the distribution of the response to the shelter usage question?

- What is the multiplicity estimate and apportionment count by age/sex group? How many persons were added by the adjustment for casual users (if we use the adjustment)?

- How many replications were omitted to account for soup kitchen respondents who also used shelters?

- What are the answers to the above questions by county and state?

- For how many counties was it not appropriate to use the multiplicity estimator?

4. **P- and E-Sample Missing Data**

There are several types of missing data known to affect the Accuracy and Coverage Evaluation. These include noninterviewed households, item missing data, and unknown status for variables such as residence, match, and correct enumeration.

To compensate for missing data we have 3 procedures in place. First, a noninterview adjustment that compensates households that could not be reached for an A.C.E. interview. Second, a characteristic imputation to fill in values for person characteristics that are missing. Finally, an procedure to estimate the probability of match, residence, or correct enumeration for those persons for whom we do not have an exact figure.

Questions to answer:

- Level and degree of household nonresponse? by State? by ACERO?
- Distribution of sample sizes and adjustment factors by noninterview adjustment cell?
- Summary of sample sizes and estimated probabilities for match, enumeration and residence cells?
- Summary at the post-stratum level of effects of each individual missing data adjustment.
- Level and degree of item missing data? by State? by ACERO?
- Level and degree of missing status? by State? by ACERO?
- Distributions of imputed and non-imputed characteristics by imputation categories?
- Number and percent imputed for certain race and ethnic characteristics?

5. **Dual System Estimation (DSE)**

DSE was used for the 1990 Census Post-Enumeration survey, the 1995 and 1996 Census Tests, and the Census 2000 Dress Rehearsal. For the Dress Rehearsal, DSE was used in conjunction with raking. For Census 2000 we will use DSE for the Accuracy and Coverage Evaluation (A.C.E.); but there will be no raking.

We will use DSE as part of the Census 2000 A.C.E., to adjust for the undercount for the redistricting file. DSE will not be used for the apportionment counts.

Questions to answer:

- For each A.C.E. post-stratum what are the weighted and unweighted components of the DSE (census count, insufficient information (II), E-sample total, correct enumerations, non-mover matches and total non-movers, out-mover matches and total out-movers, total in-movers)? What are these summary statistics after collapsing age/sex? For race/origin groups? Owners and renters? High and low mail return rate? MSA/TEA group?

- What is the coverage correction factor, undercount rate and undercount by post-stratum?

- Was any collapsing of the 448 post-strata necessary due to sample size? What was the sample size (P-sample and E-sample) by post-strata?

- What is the consistency between the E-sample and P-sample responses for each of the post-stratification variables for not imputed and total persons?

6. A.C.E. Variance Estimation

For the 2000 Census, the methodology that will be used to calculate the variances will be a stratified jacknife. This methodology will reflect the double sampling, TES, missing data and DSE.

The variance estimation summaries will focus on the different components of the variance estimate. We will conduct a quantitative analysis of the contribution of each variance component (such as the variance that results from imputation of missing data, i.e. , correct enumeration probability and p-sample matching probability) to the overall variance estimate of the dual system estimate.

Questions to answer:

- How do observed coefficients of variation compare to expected coefficients of variation for selected demographic groups? by post-stratum? By State? By Congressional District?

- What's the contribution of imputing match probability to the overall variance estimates? By post-strata?

- How do the erroneous enumerations contribute to the overall variance estimate? By post-strata?

- C How do the various stages of sampling contribute to the overall variance estimate? By post-strata?

7. **Long Form Estimation/ Variance Estimation**

Long form was used in the1990 Census. However, the long form was not used in 1995 and 1996 Census tests. The long form was used in Census 2000 dress rehearsal. For dress rehearsal estimation is used as an operational test and will not produce official long form data estimates. The long form estimation will use a weighting approach including a raking methodology for the dress rehearsal.

Long form estimation for Census 2000 will use a weighting approach which will utilize raking methodology. Raking will use marginal controls from the census data which are not corrected for coverage error. Then coverage correction factors will be applied at the person level to the results of raking.

We will review operational tallies which provide on overview of the results from implementation of the weighting and variance estimation.

<u>Questions to answer:</u>

- C What are the estimates of householders before applying coverage correction factors?
- C What estimates of householders after applying coverage correction factors?
- C What are actual variances for long form estimates at different levels of geography?
- C What are the generalized variances for long form entities at different level of geography?

8. **A.C.E. Weight Trimming**

Due to large block and small block cluster subsampling and oversampling of difficult to enumerate blocks and inconsistent blocks, there will be variation in the unbiased weights for the A.C.E. sample design. If this variation is too large, some weight trimming may be implemented. For the Dress Rehearsal, a small amount of weight truncation was implemented with no increase in other weights to account for the truncation (not necessary for dual system estimation). The A.C.E. design is not finalized. However, some weight trimming (or truncation) may be needed. Research continues to determine how to truncate weights if needed.

<u>Questions to answer:</u>

- C How do the dual system estimates and variances compare using trimmed weights and not trimming the weights?

- What are the weight distributions by trimming Cells before and after trimming?
- What are the weighted non-matches and non-match rates before and after trimming by trimming Cells? and by Region/MSA/TEA areas?
- What are the weighted erroneous enumerations and erroneous enumeration rates before and after trimming by trimming Cells? and by Region/MSA/TEA areas?
- What are the weighted E-sample and P-sample person estimates before and after trimming by State?
- What are the E-sample and P-sample American Indians on AIR estimates before and after trimming.

9. Generalized Variances

For small geographic areas such as blocks, direct variance estimates are smoothed by fitting a standard GATT curve. The parameters from these models are provided to data users for calculation of standard errors. The methodology is equivalent to what was used in dress rehearsal.

The generalized variance summaries for A.C.E. data will focus on how well the generalized variance function model approximated the estimated variances at various geographic levels and characteristic estimates.

The generalized variance research for Long Form will focus on use of the design factor to approximate the variances of sample estimates at various levels of geography, particularly census tracts and block groups.

Questions to answer:

- How well the fit of the weighted GATT Curve model approximate the estimated variances at all levels of geography?

- Which level of geography of the data, that is fit into the model, estimates the other levels of geography the best?

- How well the generalized design factors approximate the estimated variances?

- Differences in the design factors for metropolitan statistical area (MSA) versus non MSA?

- Summary of effect of removal of outliers?

10. **Block Level Estimation**

Block level estimation for the 1990 Census Post-Enumeration survey and the Census 2000 Dress Rehearsal assumed that within poststrata the estimated coverage factor applied to all small areas (synthetic assumption). Controlled rounding was used to create whole person records in order to correct for undercount or overcount.

We will use the synthetic assumption and controlled rounding for block level estimation for the Census 2000 Accuracy and Coverage Evaluation (A.C.E.).

Questions to answer:

- What are the rounded and unrounded counts by post-stratum for each block (prepare a block file)?

- Using this file what are the rounded and unrounded counts by post-stratum at the county, state, and national level?

- At the county, state, post-stratum, and national level what are the number of undercount or over count persons records created by synthetic estimation?

- What is the distribution of the relative and absolute effect of synthetic estimation on block totals?

11. **Unclassified Estimation**

Unclassified units are housing units with unknown status (occupied, vacant or nonexistent) and occupied housing units with unknown population count. The missing status or population count must be estimated prior to Population Division's Edit & Imputation. For the Census 2000 Dress Rehearsal unclassified estimation was part of the estimation for Nonresponse follow-up and Undeliverable as Addresses Vacant sampling. Since this sampling will not be part of Census 2000, a separate estimation for unclassified housing units is necessary.

The nearest-neighbor hot deck imputation method will be used for Unclassified Estimation for Census 2000.

Questions to answer:

- What is the distribution of classified and unclassified units by occupied, vacant , and delete after the completion of unclassified estimation? By LCO, County, and State? By donor/donee group?

- What is the average household size for classified and unclassified occupied units after the completion of unclassified estimation? By LCO, County, and State? By donor/donee group?

- What percentage of housing units are classified and unclassified prior to unclassified estimation? By LCO, County, and State? By donor/donee group?

- Based on the A.C.E. post-stratum of the head of household what are the number of occupied classified units and their average household size by post-stratum? By A.C.E. block clusters and non A.C.E. block cluster?

- Based on the A.C.E. post-stratum of the head of household what are the number of occupied unclassified units and their average household size after unclassified estimation by post-stratum?

- By A.C.E. block clusters and non A.C.E. block cluster?

- What are the number of persons in classified and unclassified (imputed by unclassified estimation) by A.C.E. post-stratum. By A.C.E. block clusters and non A.C.E. block cluster?

12. Housing Unit (HU) Dual System Estimation (DSE)

HU DSE was used to support evaluation of the 1990 Census using data from the 1990 Census Post-Enumeration Survey. For the Census 2000 Dress Rehearsal, HU DSE was planned to support housing unit long form weighting as well as for evaluation. However, HU DSE was canceled for the Dress Rehearsal.

We will use HU DSE as part of the Census 2000 Accuracy and Coverage Evaluation (A.C.E.) to support housing unit long form weighting and to evaluate HU coverage.

Questions to answer:

- For each A.C.E. housing unit post-stratum what are the weighted and unweighted components of the DSE (census count, E-sample total, correct enumerations, matches and total)? What are the resulting, erroneous enumeration, and match rates by post-stratum?

- What is the coverage correction factor, undercount rate and undercount by post-stratum?

- What is the consistency between the E-sample and P-sample responses for each of the post-stratification variables?

13. **A.C.E. Targeted Extended Search**

The Census 2000 A.C.E. will implement a targeted surrounding block search operation, hereinafter, Targeted Extended Search or simply TES. The rationale is to develop a criteria by which to identify block clusters that will benefit the most from surrounding block search. In 1990, the majority of the matches found in surrounding blocks were in addresses that were incorrectly geocoded in a surrounding block. Thus, the targeting criterion or criteria will be based on information or factors which give evidence of census geocoding error. For instance; block clusters with high rates of A.C.E. housing units without a census match (coded CI) and/or census geocoding error (coded GE) will be included in the TES. The goal of this operation is to increase both the matching and correct enumeration rates compared to a design that limits the search to the A.C.E. block cluster.

<u>Questions to be answered:</u>

- Summary of TES block clusters by certainty strata, TES sampling universe and out of scope?
- Summary of TES block clusters by State and A.C.E sampling stratum?
- What % of gecoding and P-Sample nonmatches are accounted for in the TES blocks based on initial housing unit information? Include distribution of gecoding and nonmatches for A.C.E. block clusters.
- What's the effect of the TES on the DSE for total population? By poststrata? By race/Hispanic Origin group?
- What is the effect of TES on each Post-stratum's match and correct enumeration rate?
- What's the effect of the TES operation on the reliability of the DSE's? By poststrata? By race/Hispanic Origin group?

ESCAP MEETING NO. 6 - 03/22/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 6

March 22, 2000

Prepared by: Maria Urrutia and Annette Quinlan

The sixth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on March 22, 2000 at 10:30. The agenda for the meeting was an overview of A.C.E. 2000 Evaluations.

Persons in attendance:

Kenneth Prewitt
William Barron
Nancy Potok
Paula Schneider
John Thompson
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
Susan Miskura
Tommy Wright
Donna Kostanich
Louisa Miller
Raj Singh
Rita Petroni
Carolee Bush
Maria Urrutia
Annette Quinlan

I. A.C.E. 2000 Evaluations

The A.C.E. 2000 evaluations are arranged into two basic categories: (1) operational summaries and (2) A.C.E. evaluations. These were discussed with the ESCAP.

1. Operational Summaries

These are evaluations that will provide real time results of key A.C.E. Sampling and Estimation activities. These evaluations result from the verification process, and as such will be available for the decision process on determining whether to use A.C.E. to adjust the redistricting data required by PL-94-171. Donna Kostanich distributed a brief overview of Census 2000 Operational Analyses for Sampling and Estimation that will be conducted by DSSD staff. This overview included a short description of key A.C.E. processes and included a summary of the measures that will be produced for each activity. Comments or additional suggestions are welcome and should be sent to Donna Kostanich by May 5.

The highlights of these evaluations included A.C.E. Sample Design, P- and E-Sample Missing Data, Dual System Estimation (DSE), A.C.E. Targeted Extended Search, and A.C.E. Weight Trimming. The discussion on A.C.E. weight trimming raised a sensitive issue. The potential use of weight trimming may conflict with our plans to completely prespecify the A.C.E. methodology. It was decided that the ESCAP will review and approve any usage of weight trimming for the A.C.E. This issue will be revisited at a future ESCAP meeting.

2. A.C.E. Evaluations

Ruth Ann Killion distributed study plans for three sets of A.C.E. Evaluations. They described A.C.E. Operations Evaluations, Coverage Evaluations for A.C.E., and A.C.E. Statistical Design and Estimation Evaluations. These are long term evaluations based on additional research, such as re-interview studies, that would be used to assess the overall accuracy of the A.C.E. As an attachment to the study plans, background information for the Total Error Model study evaluations was provided. All handouts are included with these minutes.

The A.C.E. Evaluations conducted by PRED staff will occur after Census and A.C.E. processing operations have been completed. The ESCAP requested that Ruth Ann Killion and her staff review the evaluations to determine whether any preliminary findings would be available to assist in the review of the A.C.E. results prior to releasing redistricting data.

There was discussion of the concept of the Total Error Model and the key evaluations that will be used to construct it. In short, the Total Error Model presents an overall quantification of the sampling and nonsampling errors associated with the A.C.E. and the Census data adjusted based on the A.C.E. results. The Total Error Model incorporates and combines the individual evaluation components to produce an overall measurement of accuracy. The results of the Total Error Model Evaluation will not be available until late 2001.

II. Next Meeting

The next meeting scheduled for Wednesday April 12, 2000 will discuss A.C.E. Post-stratification.

<u>ESCAP Committee</u>

Kenneth Prewitt
William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson, Chair
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

cc:

Teresa Angueira
Bill Bell
Debbie Bolton
Genny Burns
Carolee Bush
Gerald Gates
Ed Gore
Dave Hubble
Donna Kostanich
Ellen Lee
Charlene Leggieri
Don Malec
Betsy Martin
Catherine Miller
Fay Nash
Sally Obenski
Miguel Perez
Ed Pike
Magdalena Ramos
Gregg Robinson
Raj Singh
Maria Urrutia
Signe Wetrogen
David Whitford
Henry Woltman
Tommy Wright